

Sintassi come dinamica oscillatoria

Struttura gerarchica del linguaggio e accoppiamento di fase nelle reti neurali corticali

I. L'imbarazzo della struttura

Si direbbe che c'è un'asimmetria curiosa, e fino ad oggi forse non sufficientemente tematizzata, tra ciò che la linguistica teorica afferma della struttura sintattica e ciò che le neuroscienze cognitive riescono a mostrare del suo substrato biologico. La linguistica — nel solco della tradizione inaugurata da Chomsky e successivamente affinata attraverso il Programma Minimalista — descrive la sintassi come un sistema di operazioni formali che generano strutture gerarchiche discontinue: non sequenze lineari di parole, ma alberi di costituenti annidati, ciascuno dominante rispetto ai propri elementi interni e dominato rispetto all'unità immediatamente sovraordinata. L'intuizione fondamentale è che la mente, nel processare la frase *"Il filosofo che aveva letto Frege comprende la logica"*, non procede da sinistra a destra come un automa a stati finiti, ma costruisce simultaneamente molteplici livelli di dipendenza strutturale: la relazione soggetto-verbo principale, la proposizione relativa incassata, i sintagmi nominali con le loro teste lessicali.

Le neuroscienze, viceversa, hanno a lungo guardato al linguaggio attraverso la lente di una topografia corticale relativamente stabile: area di Broca per la produzione e l'elaborazione sintattica, area di Wernicke per la comprensione lessicale, fascicolo arcuato come via di connessione, con successive raffinazioni che hanno aggiunto regioni temporali superiori e prefrontali dorsali. Questa mappatura, pur preziosa, resta fondamentalmente *spaziale*: localizza *dove* avviene qualcosa, senza spiegare *come* processi distribuiti su aree distanti vengano integrati in rappresentazioni coerenti. Rimane silenziosa, soprattutto, sulla questione più spinosa: in che senso il cervello *possiede* struttura gerarchica, quando tutto ciò che i neuroni fanno è scambiare segnali elettrochimici distribuiti nel tempo?

La proposta che questo saggio intende articolare è la seguente: la gerarchia sintattica non è soltanto rappresentata nel cervello — come un oggetto mentale che occupa una porzione di spazio neurale — ma è *attuata* mediante la dinamica temporale delle oscillazioni corticali. Più specificamente, la struttura costituente

emerge dall'accoppiamento sistematico tra bande di frequenza diverse, dove i ritmi più lenti modulano i ritmi più rapidi in una gerarchia temporale che rispecchia la gerarchia strutturale del linguaggio. Se questa tesi è corretta, allora la sintassi non è un'astrazione teorica comodamente separabile dal suo supporto fisico: è, in un senso tecnico e non metaforico, una proprietà dell'oscillazione cerebrale.

Questo non è soltanto un risultato empirico di interesse neuroscientifico. È un fatto filosoficamente gravido di conseguenze, perché obbliga a riconsiderare la natura dello strutturale in linguistica, il rapporto tra forme logiche e processi temporali, e la stessa distinzione — cara alla tradizione chomskyana — tra competenza ed esecuzione.

II. Il Programma Minimalista e l'operazione di Merge

Per apprezzare la portata della proposta, occorre innanzitutto rendere esplicita la concezione della struttura sintattica che il Programma Minimalista di Chomsky ha elaborato a partire dagli anni Novanta del Novecento, in particolare nella versione cristallizzata in *The Minimalist Program* (1995) e nei successivi scritti sulla struttura delle frasi e sulla biolinguistica.

Il fulcro del Minimalismo è l'operazione denominata **Merge** (fusione): un'operazione binaria, ricorsiva e set-teoricamente elementare, che prende due oggetti sintattici α e β — siano essi elementi lessicali o oggetti già costruiti da precedenti applicazioni di Merge — e produce un nuovo oggetto sintattico $\{\alpha, \beta\}$. La potenza generativa dell'intero sistema linguistico discende, nella concezione minimalista, da questa singola operazione: la ricorsività di Merge è sufficiente a generare la struttura gerarchica illimitata che contraddistingue il linguaggio umano.

La frase viene dunque costruita *bottom-up*, partendo dagli elementi lessicali e procedendo verso unità sempre più grandi. Il sintagma "*legge Frege*" è il risultato di $\{\text{legge}, \text{Frege}\}$, dove *legge* è la testa verbale che seleziona *Frege* come suo complemento. Il sintagma "*aveva letto Frege*" è $\{\text{aveva}, \{\text{letto}, \text{Frege}\}\}$, e così via fino alla derivazione dell'intera frase. Ogni nodo interno dell'albero sintattico è un oggetto generato da un'applicazione di Merge; ogni ramificazione è binaria; ogni struttura è, per costruzione, gerarchica e non lineare.

Due aspetti di questa concezione meritano particolare sottolineatura per i fini del presente argomento.

Il primo riguarda la **linearizzazione**. Il linguaggio, nella sua forma fonologica, è necessariamente lineare: le parole si susseguono nel tempo. Ma la struttura sintattica generata da Merge è fondamentalmente non lineare, priva di ordine intrinseco. L'algoritmo che mappa la struttura gerarchica sulla sequenza

temporale lineare — quello che Chomsky e collaboratori chiamano *linearizzazione* e che Kayne ha formalizzato attraverso il *Linear Correspondence Axiom* — è logicamente posteriore alla costruzione sintattica. Ne consegue che la linearità è un vincolo imposto dall'interfaccia fonologica su una struttura che è, nella sua essenza, *atemporale* o, meglio, *sovra-temporale*.

Il secondo aspetto riguarda la **computazione sintattica come sistema chiuso**. Nella concezione minimalista pura, la sintassi è un sistema autonomo che opera su oggetti formali e produce strutture valutate alle interfacce — la Forma Fonetica (PF) e la Forma Logica (LF) — ma che al proprio interno è indifferente sia al significato lessicale che alla struttura prosodica. La celebre modularità della sintassi chomskyana è una tesi sull'incapsulamento informativo: il sistema sintattico non *vede* il significato mentre computa la struttura.

Queste due proprietà — la non-linearità intrinseca della struttura e l'autonomia computazionale della sintassi — pongono un problema filosoficamente acuto al momento in cui si tenta di dare a tale concezione un'implementazione neurale plausibile. Il cervello è un sistema dinamico che opera in tempo reale; i suoi stati si dispiegano necessariamente in sequenza temporale; non possiede, almeno in senso ovvio, nulla che assomigli a strutture set-teoriche atemporalmente. Come può, allora, implementare un'operazione come Merge, che per definizione opera su insiemi non ordinati?

La risposta che le neuroscienze computazionali cominciano a elaborare è paradossale nella sua eleganza: il cervello *non* implementa Merge come operazione su oggetti atemporalmente. Piuttosto, usa il *tempo stesso* come dimensione strutturale, codificando la gerarchia non nell'asse spaziale della corteccia ma nell'asse temporale delle sue oscillazioni. La struttura, in questo senso, non è contenuta nel cervello: è il cervello che oscilla.

III. Le oscillazioni corticali come substrato computazionale

Prima di procedere alla trattazione del loro ruolo specifico nel linguaggio, è necessario richiamare alcuni elementi fondamentali della fisiologia delle oscillazioni corticali, senza i quali la proposta successiva resterebbe priva del necessario ancoraggio empirico.

Le oscillazioni neurali — fluttuazioni ritmiche del potenziale di campo locale (LFP, *local field potential*) e dell'attività elettroencefalografica — non sono artefatti del rumore biologico ma riflettono la sincronizzazione di popolazioni neuronali che si trovano in stati di eccitazione e inibizione ciclicamente alternati. La struttura di queste oscillazioni è gerarchica per natura: le bande di frequenza convenzionalmente distinte (delta, theta, alfa, beta, gamma) non sono categorie arbitrarie ma corrispondono a meccanismi biofisici distinti e a scale temporali di integrazione corticale differenti.

La banda **delta** (circa 0.5–4 Hz) riflette oscillazioni su scale temporali dell'ordine di centinaia di millisecondi: i cicli delta corrispondono, nel dominio linguistico, alla durata approssimativa delle frasi e delle clausole. La banda **theta** (circa 4–8 Hz) opera sulla scala temporale della sillaba e della parola: un ciclo theta dura approssimativamente 125–250 millisecondi, e questa durata si sovrappone sorprendentemente alla durata media della sillaba nelle lingue del mondo. La banda **beta** (circa 12–30 Hz) è associata al mantenimento di rappresentazioni in memoria di lavoro e alla segnalazione di stati stabili, con un ruolo particolare nei processi di integrazione sintattica a livello della parola e del morfema. La banda **gamma** (circa 30–80 Hz, e oltre) riflette l'elaborazione locale ad alta risoluzione, correlata con il riconoscimento di tratti fonologici e la binding di rappresentazioni elementari.

Questa stratificazione per frequenze non è casuale: riflette la struttura geometrica e biofisica dei circuiti corticali. Le oscillazioni lente emergono da circuiti cortico-talamici a lungo raggio, che integrano informazioni distribuite su vaste regioni; le oscillazioni rapide nascono da circuiti locali corticali, in particolare dall'interazione tra interneuroni inibitori (soprattutto le cellule parvalbuminergiche a canestro) e neuroni piramidali eccitatori. La dualità eccitazione-inibizione è, al fondo, il motore biofisico di ogni oscillazione neurale: il ritmo emerge dall'equilibrio instabile tra l'eccitazione che accende le popolazioni e l'inibizione che le silenzia, per poi lasciarle nuovamente eccitare.

Ciò che rende questo sistema potenzialmente rilevante per il linguaggio è la sua proprietà di **accoppiamento cross-frequenza** (*cross-frequency coupling*, CFC): le oscillazioni di bande diverse non sono indipendenti ma si influenzano reciprocamente in modi sistematici. Il fenomeno più studiato e teoricamente più rilevante è l'**accoppiamento fase-ampiezza** (*phase-amplitude coupling*, PAC): la fase di un'oscillazione lenta modula l'ampiezza di un'oscillazione rapida. In termini concreti, questo significa che le creste e i valli dell'oscillazione theta determinano quando l'oscillazione gamma è forte o debole; oppure che il delta organizza la struttura interna delle oscillazioni theta. Il risultato è una gerarchia temporale: il ritmo più lento funge da impalcatura temporale entro cui i ritmi più rapidi svolgono il loro lavoro computazionale locale.

È in questa gerarchia temporale che, come si argomenterà, trova la propria controparte neurale la gerarchia strutturale della sintassi.

IV. La convergenza: Giraud, Poeppel e la teoria oscillatoria del parsing linguistico

La proposta che le oscillazioni corticali svolgano un ruolo non accessorio ma costitutivo nell'analisi sintattica del linguaggio deve molto al lavoro pionieristico di Anne-Lise Giraud e David Poeppel, in particolare all'articolo seminale del 2012 su *Nature Neuroscience*, intitolato "*Cortical oscillations and speech processing*:"

emerging computational principles and operations". Sebbene il lavoro di Giraud e Poeppel si concentrasse primariamente sul livello fonologico e sillabico, le sue implicazioni si estendono in modo naturale e fecondo al livello sintattico superiore.

La tesi di Giraud e Poeppel è che la corteccia uditiva si accoppia in modo selettivo al segnale acustico del parlato attraverso un meccanismo di *entrainment* oscillatorio: le oscillazioni corticali endogene si sincronizzano con le fluttuazioni temporali dell'involuppo acustico del parlato, a diverse scale temporali. Le oscillazioni theta (4–8 Hz) si sincronizzano con la struttura sillabica, le cui fluttuazioni di ampiezza e spettro si dispiegano appunto alla frequenza di 4–8 eventi al secondo nelle lingue del mondo. Le oscillazioni gamma, intanto, operano all'interno di ogni ciclo theta per analizzare la struttura fonemica interna alla sillaba.

Questo modello, nella sua versione originale, propone dunque una gerarchia a due livelli: theta per la sillaba, gamma per il fonema. Ma il linguaggio naturale possiede una gerarchia strutturale che si estende ben al di là del livello sillabico: parole, morfemi, sintagmi, clausole, frasi. La domanda naturale — e teoricamente urgente — è: questa gerarchia si prolunga verso l'alto, con bande di frequenza ancora più basse che corrispondono a livelli strutturali ancora più ampi?

Le evidenze sperimentali accumulate nel decennio successivo suggeriscono che la risposta sia affermativa. Studi di EEG e MEG (*magnetoencefalografia*) hanno mostrato che le oscillazioni delta (1–4 Hz) si sincronizzano con la struttura prosodica e con i confini sintattici di frase: i cicli delta corrispondono, nella loro durata, alla tipica durata delle frasi nel parlato continuo. Studi di elettrocorticografia (ECoG) su pazienti neurochirurgici — che permettono una risoluzione spazio-temporale molto superiore rispetto all'EEG di superficie — hanno confermato la presenza di accoppiamento sistematico tra delta e theta, e tra theta e gamma, durante l'ascolto di parlato naturale e l'elaborazione di frasi con diversa struttura sintattica.

Il quadro emergente è quello di una **gerarchia oscillatoria corticale** che rispecchia punto per punto la gerarchia strutturale del linguaggio:

- Gamma (30–80 Hz): livello dei tratti fonologici e dei segmenti
- Theta (4–8 Hz): livello della sillaba e del morfema
- Beta (12–30 Hz): livello della parola e dell'integrazione lessicale
- Delta (0.5–4 Hz): livello del sintagma, della clausola e della frase

L'accoppiamento fase-ampiezza tra questi livelli non è casuale: è la firma elettrofisiologica di un meccanismo computazionale mediante il quale l'elaborazione locale (ad alta frequenza, ad alta risoluzione) viene organizzata e temporizzata dall'impalcatura globale (a bassa frequenza, a bassa risoluzione). Il tempo, in questo senso, diventa la dimensione in cui la struttura gerarchica si iscrive.

Rimane tuttavia da spiegare, con il rigore formale che l'argomento esige, come questo meccanismo funzioni in termini matematici precisi. È a questo livello che la proposta diventa pienamente articolabile — e pienamente falsificabile.

V. L'accoppiamento fase-ampiezza: architettura matematica

La nozione di accoppiamento fase-ampiezza richiede una trattazione matematica rigorosa per non ridursi a una metafora evocativa priva di contenuto empiricamente testabile. Il punto di partenza è la decomposizione di un segnale neurale — tipicamente il potenziale di campo locale o il segnale EEG — nelle sue componenti oscillatorie a diverse frequenze.

Dato un segnale $x(t)$, la sua trasformata di Hilbert $\hat{x}(t)$ è definita come:

$$\hat{x}(t) = (1/\pi) \cdot \text{P.V.} \int_{-\infty}^{+\infty} x(\tau)/(t - \tau) d\tau$$

dove P.V. denota il valore principale di Cauchy. Il segnale analitico $z(t)$ è il segnale complesso costruito come:

$$z(t) = x(t) + i \cdot \hat{x}(t) = A(t) \cdot e^{i\varphi(t)}$$

dove $A(t) = |z(t)| = \sqrt{x(t)^2 + \hat{x}(t)^2}$ è l'involuppo di ampiezza istantanea e $\varphi(t) = \arctan(\hat{x}(t)/x(t))$ è la fase istantanea. Questa decomposizione è non-ambigua a condizione che il segnale sia a banda stretta in prossimità della frequenza di interesse, condizione garantita dal filtraggio a banda passante applicato prima del calcolo della trasformata di Hilbert.

Per misurare l'accoppiamento fase-ampiezza tra una banda di frequenza lenta f_L (ad esempio theta, 4–8 Hz) e una banda rapida f_H (ad esempio gamma, 30–80 Hz), si procede nel modo seguente. Si estraggono mediante filtraggio a banda le componenti rispettive del segnale:

$$\begin{aligned} x_L(t) &\rightarrow \varphi_L(t) \quad (\text{fase dell'oscillazione lenta}) \\ x_H(t) &\rightarrow A_H(t) \quad (\text{ampiezza dell'oscillazione rapida}) \end{aligned}$$

Il **Modulation Index** (MI), introdotto da Tort e collaboratori nel 2010, è la misura di accoppiamento più ampiamente adottata. Si divide il ciclo della fase lenta in N bin di uguale ampiezza (tipicamente $N = 18$, corrispondenti ad archi di 20° ciascuno), e per ciascun bin k si calcola l'ampiezza media dell'oscillazione rapida:

$$\langle A_H \rangle_k = (1/|T_k|) \cdot \int_{T_k} A_H(t) dt$$

dove T_k è l'insieme degli istanti in cui $\varphi_L(t)$ cade nel k -esimo bin. Si costruisce quindi la distribuzione discreta di probabilità:

$$p(k) = \langle A_H \rangle_k / \sum_{j=1}^N \langle A_H \rangle_j$$

Il Modulation Index è la divergenza di Kullback-Leibler tra questa distribuzione p e la distribuzione uniforme U (che rappresenta l'assenza di accoppiamento), normalizzata dal logaritmo di N per renderla invariante rispetto alla scelta del numero di bin:

$$MI = D_{\{KL\}}(p \parallel U) / \log(N) = [\sum_{k=1}^N p(k) \cdot \log(p(k)/U(k))] / \log(N)$$

dove $U(k) = 1/N$ per ogni k . Si noti che $D_{\{KL\}}(p \parallel U) \geq 0$ per la disuguaglianza di Gibbs, con uguaglianza se e solo se $p = U$. Dunque $MI \in [0, 1]$, con $MI = 0$ in assenza di qualunque accoppiamento (distribuzione uniforme dell'ampiezza gamma rispetto alla fase theta) e $MI \rightarrow 1$ in presenza di accoppiamento massimo (tutta l'energia gamma concentrata in un singolo bin di fase).

Una misura alternativa, concettualmente correlata ma formalmente distinta, è il **Phase-Locking Value** (PLV), che quantifica la sincronia di fase tra due segnali a frequenze diverse — ad esempio la fase dell'oscillazione theta in una regione corticale e la fase dell'oscillazione gamma in una regione adiacente:

$$PLV = | \langle e^{i(\varphi_H(t) - n \cdot \varphi_L(t))} \rangle_t |$$

dove n è il rapporto intero di frequenze (detto *n:m phase locking ratio*, con $m = 1$ per il caso standard) e le parentesi angolari denotano la media temporale. Il PLV è un numero reale nell'intervallo $[0, 1]$: vale 0 quando le differenze di fase sono uniformemente distribuite nel tempo (nessuna sincronia) e 1 quando la differenza di fase è costante nel tempo (sincronia perfetta).

Questi due indici — MI e PLV — misurano aspetti complementari dell'accoppiamento cross-frequenza: il primo la modulazione dell'ampiezza di un'oscillazione rapida da parte della fase di una lenta; il secondo la coerenza di fase tra oscillazioni a frequenze diverse. Entrambi risultano significativamente elevati nelle regioni corticali temporali superiori e nelle regioni prefrontali inferiori durante l'elaborazione di frasi con struttura sintattica complessa, e si riducono in modo corrispondente in presenza di frasi con struttura degradata o di sequenze di parole prive di struttura (liste di parole non correlate, spesso usate come condizione di controllo negli studi neuroimaging del linguaggio).

VI. Dinamica dei sistemi accoppiati: il modello di Kuramoto

Per comprendere in che modo questa struttura di accoppiamento emerga — non come proprietà imposta dall'esterno ma come attrazione di uno stato del sistema dinamico corticale — è utile ricorrere al modello di Kuramoto, che fornisce l'architettura matematica più trasparente per la sincronizzazione di oscillatori accoppiati.

Si consideri una popolazione di N oscillatori, ciascuno con una frequenza naturale ω_i e uno stato descritto dalla sua fase istantanea $\theta_i(t)$. In assenza di accoppiamento, ciascun oscillatore evolve linearmente: $d\theta_i/dt = \omega_i$. Con accoppiamento globale di forza K , le equazioni del moto diventano:

$$d\theta_i/dt = \omega_i + (K/N) \cdot \sum_{j=1}^N \sin(\theta_j - \theta_i)$$

Il parametro d'ordine del sistema, che misura il grado di sincronizzazione collettiva, è il numero complesso:

$$r(t) \cdot e^{i\psi(t)} = (1/N) \cdot \sum_{j=1}^N e^{i\theta_j(t)}$$

dove $r \in [0, 1]$ è il modulo (il grado di sincronizzazione) e ψ è la fase media della popolazione. In termini di r e ψ , le equazioni del moto si riscrivono elegantemente come:

$$d\theta_i/dt = \omega_i + K \cdot r \cdot \sin(\psi - \theta_i)$$

Questa forma rivela la struttura del fenomeno: ogni oscillatore è attratto verso la fase media ψ con una forza proporzionale al prodotto $K \cdot r$ — la forza di accoppiamento per il grado di coerenza già raggiunto. Il sistema è autoamplificante: man mano che la coerenza aumenta, la forza di attrazione cresce, reclutando oscillatori sempre più distanti dalla fase media. L'analisi nel limite termodinamico $N \rightarrow \infty$ (con la distribuzione delle frequenze naturali $g(\omega)$ lorentziana di larghezza γ e centro ω_0) mostra una transizione di fase del secondo ordine: per $K < K_c = 2\gamma$ il sistema rimane incoerente ($r = 0$ all'equilibrio), mentre per $K > K_c$ si rompe la simmetria e la soluzione stazionaria ha:

$$r = \sqrt{1 - K_c/K}$$

Questa transizione è la formalizzazione matematica del concetto intuitivo di "sincronia corticale": la corteccia non è mai completamente incoerente né completamente sincrona, ma si colloca in un regime vicino alla transizione critica ($K \approx K_c$), un regime chiamato in letteratura **criticità** o **edge of chaos**, nel quale la trasmissione di informazione e la flessibilità computazionale sono massimizzate.

Per modellare la gerarchia oscillatoria del linguaggio, il modello di Kuramoto va esteso in almeno due direzioni. La prima è l'introduzione di **accoppiamento gerarchico**: oscillatori a frequenza diversa non sono accoppiati tra loro in modo omogeneo, ma attraverso una struttura a livelli, in cui gli oscillatori del livello superiore (più lento) modulano la frequenza degli oscillatori del livello inferiore (più rapido). Se si denota con $\Theta(t)$ la fase dell'oscillatore lento e con $\theta_i(t)$ le fasi degli oscillatori rapidi, un modello minimale di accoppiamento gerarchico assume la forma:

$$d\theta_i/dt = \omega_i^{\{(fast)\}} + K_{\{ff\}} \cdot r^{\{(fast)\}} \cdot \sin(\psi^{\{(fast)\}} - \theta_i) + K_{\{top-down\}} \cdot \sin(n \cdot \Theta - \theta_i)$$

dove il terzo termine rappresenta l'influenza discendente dell'oscillatore lento: la fase dell'oscillazione rapida θ_i tende a sincronizzarsi con l' n -esimo armonico della fase lenta Θ . Questo termine di accoppiamento discendente (*top-down*) è il correlato dinamico del PAC: quando $K_{\{top-down\}}$ è sufficientemente grande, l'ampiezza media dell'oscillazione rapida — proporzionale a $r^{\{(fast)\}}$ — risulta massima in un particolare intervallo di fase di Θ , esattamente come il Modulation Index misura sperimentalmente.

La seconda estensione rilevante è l'introduzione di **ritardi di trasmissione** (*transmission delays*): i segnali neurali non si propagano istantaneamente, ma con ritardi dell'ordine di decine di millisecondi a causa della velocità finita di conduzione degli assoni. L'equazione di Kuramoto con ritardi è:

$$d\theta_i/dt = \omega_i + (K/N) \cdot \sum_{j=1}^N \sin(\theta_j(t - \tau_{ij}) - \theta_i(t))$$

dove τ_{ij} è il ritardo di trasmissione tra l'oscillatore j e l'oscillatore i . I ritardi di trasmissione introducono una ricchezza fenomenologica considerevole: possono stabilizzare stati di sincronizzazione sfasata (*cluster states*), indurre dinamiche caotiche, o generare stati di sincronizzazione multistabile. Nel contesto corticale, si ritiene che i ritardi anatomicamente determinati contribuiscano a definire le frequenze di risonanza dei circuiti a lungo raggio, stabilendo così una corrispondenza tra la geometria della connettività anatomica e lo spettro delle frequenze dominanti nell'attività spontanea.

VII. Verso una fenomenologia dell'oscillazione: il tempo come struttura

Il formalismo delineato nelle sezioni precedenti può essere letto come una proposta puramente computazionale — una descrizione di meccanismi biofisici senza implicazioni filosofiche sostanziali. Sarebbe però un errore arrestarsi a questa lettura riduzionista, che lascia senza risposta la domanda più profonda: che rapporto ha questa dinamica con l'*esperienza* del linguaggio, con la

comprensione che il parlante-ascoltatore ha della struttura della frase che elabora?

Questa domanda rinvia direttamente al problema della coscienza linguistica e della sua relazione con i suoi substrati fisici — una delle questioni centrali della filosofia della mente contemporanea. Ma vi è un aspetto più specifico e, per certi versi, ancora più urgente: il rapporto tra la temporalità dell'oscillazione e la temporalità dell'esperienza linguistica.

Husserl, nelle *Lezioni sulla fenomenologia della coscienza interna del tempo* (1928), aveva mostrato che l'esperienza di un oggetto temporalmente esteso — una melodia, una frase pronunciata — non è possibile attraverso una serie di percezioni istantanee puntiformi. Ogni *adesso* (Jetzt-Punkt) porta con sé una *ritenzione* delle fasi appena trascorse e una *protenzione* di ciò che sta per venire: queste due dimensioni — il passato ancora presente come modificazione interna dell'ora vivente, e il futuro anticipato come vuoto intenzionale — sono costitutive dell'esperienza stessa, non aggiunte secondarie ad atti discreti di percezione.

La ritenzione husserliana non è un ricordo: è l'inerzia temporale dell'esperienza, la sua capacità di tenere insieme ciò che è appena stato con ciò che è. Nella percezione di una frase, questa struttura tripartita (ritenzione, adesso, protenzione) opera a scale temporali diverse in modo simultaneo: la ritenzione del fonema appena udito si sovrappone alla ritenzione del morfema, della parola, del sintagma — e le protenzioni anticipano, a scale corrispondenti, i completamenti attesi a ciascun livello.

Ciò che le neuroscienze oscillatorie suggeriscono è che questa struttura tripartita husserliana non è soltanto una descrizione fenomenologica dell'esperienza temporale, ma potrebbe avere un correlato funzionale preciso nell'architettura dell'accoppiamento cross-frequenza. La fase dell'oscillazione lenta (delta-theta) costituirebbe l'impalcatura della protenzione: essa anticipa la struttura temporale del segnale in arrivo, preparando le popolazioni neurali ad elaborare l'informazione nel momento in cui essa giungerà. L'oscillazione rapida (gamma), che opera localmente all'interno di ciascun ciclo lento, realizzerebbe invece il momento dell'adesso percettivo. La codifica progressiva dell'informazione nella memoria di lavoro — facilitata dal beta — corrisponderebbe alla ritenzione, al mantenimento attivo di ciò che è stato elaborato nei cicli precedenti.

Questa corrispondenza non è una prova di identità tra fenomeni fenomenologici e fenomeni neurali. Sarebbe ingenuo pensarlo, e il dualismo metodologico che separa i livelli di descrizione rimane un imperativo epistemologico ineludibile. Tuttavia, la corrispondenza strutturale è abbastanza precisa da non poter essere liquidata come coincidenza: suggerisce che la struttura temporale husserliana dell'esperienza non è un'invenzione speculativa ma cattura qualcosa di reale nell'architettura del processamento cognitivo — qualcosa che le neuroscienze oscillatorie cominciano a rendere accessibile all'indagine empirica.

VIII. La realtà della sintassi: un problema ontologico

Si può ora tornare al problema filosofico più profondo sollevato nell'apertura: la sintassi è reale a livello neurale, o è soltanto un'astrazione teorica commoda che il linguista applica ai dati senza impegnarsi nell'assunzione che essa esista come struttura *nel* parlante?

Questa questione va distinta con cura da altre che le assomigliano superficialmente. Non è la questione se i concetti della linguistica teorica siano empiricamente adeguati a descrivere il comportamento linguistico: su questo punto non vi è dubbio che lo siano, almeno parzialmente. Non è nemmeno la questione se la struttura sintattica *influenzi* i processi neurali: anche su questo vi è abbondante evidenza. La questione è più radicale: se la struttura sintattica sia ontologicamente riducibile a quei processi neurali — se il sintagma nominale *sia* una certa configurazione di accoppiamento oscillatorio — o se esista a un livello di analisi irriducibilmente autonomo.

La proposta oscillatoria, nella sua versione più ambiziosa, suggerisce una forma di **riduzionismo strutturale debole**: la gerarchia sintattica non è identica alla gerarchia oscillatoria, ma è realizzata da essa nel senso in cui una funzione matematica è realizzata da un insieme di valori. La struttura del sintagma nominale non è l'accoppiamento tra delta e theta nella corteccia temporale sinistra: è la struttura astratta di dipendenze che quell'accoppiamento implementa. Il sintagma nominale potrebbe, in linea di principio, essere implementato da un meccanismo diverso — in un cervello non biologico, o in un sistema artificiale — e resterebbe nondimeno un sintagma nominale.

Questa posizione è filosoficamente analoga al **funzionalismo** in filosofia della mente: i tipi mentali sono definiti dai loro ruoli funzionali, non dalla loro realizzazione fisica specifica. Ma il funzionalismo applicato alla sintassi incontra un ostacolo peculiare: a differenza degli stati mentali, i tipi sintattici non sono caratterizzati dal loro ruolo *causale* (i loro effetti comportamentali) ma dalla loro posizione in una struttura formale. Un sintagma nominale è tale perché partecipa a certe relazioni strutturali — dipendenza da una testa verbale, accordo con un determinante, possibile promozione a soggetto in certe costruzioni — non perché produca certi effetti.

Si può tuttavia sostenere che questa caratterizzazione strutturale-formale non sia alternativa a quella funzionale ma ne sia una versione raffinata: le relazioni strutturali che definiscono il sintagma nominale sono precisamente le relazioni che mediano certi processi cognitivi — il recupero dell'antecedente in una relazione di controllo, la computazione dell'accordo, la risoluzione di pronomi anafori. Il sintagma nominale come tipo strutturale astratto è reale in virtù del suo ruolo nell'organizzare le transizioni di stato del sistema cognitivo — e quelle

transizioni, nella proposta oscillatoria, sono modellate dall'architettura dell'accoppiamento cross-frequenza.

La conseguenza più radicale di questa posizione è che la distinzione chomskyana tra **competenza** (la conoscenza astratta della grammatica) ed **esecuzione** (il processamento linguistico in tempo reale) risulta considerevolmente indebolita. Se la struttura gerarchica è attuata dalla dinamica oscillatoria in tempo reale, allora non vi è un livello separato di competenza grammaticale che si applica, dall'esterno, al processo di elaborazione della frase: la competenza è il sistema dinamico, caratterizzato dalla propria architettura di accoppiamento e dai propri attrattori. Non si tratta di abolire la distinzione, ma di reinterpretarla: anziché distinguere tra la struttura (atemporale, formale) e il processo (temporale, fisico), la proposta oscillatoria suggerisce che la struttura *sia* un invariante del processo — la proprietà del sistema dinamico che rimane stabile attraverso le variazioni di istanza specifica.

Questo riorientamento ha una precedente illustre nella tradizione fenomenologica, benché essa non avesse gli strumenti concettuali per formularlo in modo tecnicamente preciso. Merleau-Ponty, nella *Fenomenologia della percezione* (1945), aveva sostenuto che lo schema corporeo non è una rappresentazione del corpo conservata in una memoria separata dall'azione motoria, ma si realizza nell'atto stesso del movimento: il corpo *sa* come camminare non perché consulti una mappa interna delle proprie possibilità motorie, ma perché le proprie strutture neuromuscolari sono già organizzate in modo tale da rendere possibile il cammino. Il sapere è immanente alla struttura, non trascritto in essa dall'esterno.

Analogamente, la proposta oscillatoria suggerisce che la competenza sintattica non sia una struttura di conoscenza separata dal processo di elaborazione, ma sia immanente all'architettura oscillatoria del sistema corticale: il parlante *sa* la grammatica perché le sue reti corticali sono organizzate in modo tale da realizzare, nell'elaborazione di ogni frase, la gerarchia di accoppiamento che corrisponde alla struttura gerarchica della sintassi. La grammatica è, in questo senso, una proprietà del sistema dinamico — una struttura attrattrice nello spazio degli stati del cervello oscillante.

IX. Aperture aperte: limiti e prospettive

Sarebbe disonesto concludere senza riconoscere i limiti e le difficoltà che la proposta oscillatoria deve ancora affrontare per diventare una teoria compiuta piuttosto che un programma di ricerca promettente.

Il primo e più urgente riguarda la **specificità sintattica** del segnale oscillatorio. Gli studi che mostrano accoppiamento fase-ampiezza durante l'elaborazione linguistica non dimostrano, nella maggior parte dei casi, che tale accoppiamento

sia specificamente sensibile alla struttura sintattica piuttosto che alla struttura prosodica, ritmica o semantica. Isolare il contributo della sintassi dal contributo della prosodia — che condivide molte delle scale temporali rilevanti — è metodologicamente arduo e richiede paradigmi sperimentali sofisticati, come le frasi con struttura sintattica e prosodica deliberatamente dissociate.

Il secondo limite riguarda la **direzione della causalità**. L'accoppiamento phase-ampiezza osservato durante l'elaborazione linguistica potrebbe non essere un meccanismo causale del parsing sintattico, ma un epifenomeno di altri processi cognitivi — attenzione, memoria di lavoro, anticipazione prosodica — che accompagnano l'elaborazione sintattica senza essere identici ad essa. Distinguere la causa dall'effetto richiede metodi di interferenza causale — stimolazione magnetica transcranica (TMS), optogenetica negli animali, o analisi di Granger-causality sui segnali elettrofisiologici — che restano difficilmente applicabili all'elaborazione sintattica nell'uomo.

Il terzo problema è forse il più profondo sul piano teorico: il rapporto tra la **discrezionalità della struttura sintattica** e la **continuità della dinamica oscillatoria**. La struttura sintattica è discreta: vi è o non vi è una relazione di accordo tra soggetto e verbo; una parola o è la testa del sintagma o non lo è. La dinamica oscillatoria è invece continua: le fasi e le ampiezze variano senza soluzione di continuità. Come si mappa una struttura discreta su una dinamica continua?

Una risposta possibile è che la discretezza sia una proprietà emergente degli attrattori del sistema dinamico: un sistema con attrattori multipli ben separati può esibire comportamenti discreti anche se le sue equazioni del moto sono continue. Ma dimostrare che il sistema oscillatorio corticale possiede attrattori che corrispondono strutturalmente alle categorie sintattiche discrete è un programma di ricerca la cui realizzazione è ancora in larga misura davanti a noi.

Queste difficoltà non falsificano la proposta: la rendono più interessante. Un'ipotesi che non incontrasse ostacoli non meriterebbe il nome di teoria. Ciò che la proposta oscillatoria offre, nella sua forma attuale, è qualcosa di filosoficamente e scientificamente raro: un punto di convergenza tra la teoria formale della struttura sintattica, la fisiologia delle reti neurali corticali, e la fenomenologia dell'esperienza temporale. Questo triangolo di prospettive, ciascuna irriducibile alle altre e ciascuna illuminata dalle altre, delinea uno spazio concettuale in cui alcune delle domande più antiche sulla natura del linguaggio — cosa significa per una struttura astratta essere reale, come la forma si iscrive nella materia, come il tempo porta il senso — possono essere affrontate con strumenti che Frege, Saussure e Husserl non avrebbero potuto immaginare, ma che non tradiscono lo spirito della loro interrogazione.

X. Il problema degli attrattori discreti in sistemi continui

La tensione tra la discretezza categoriale della struttura sintattica e la continuità della dinamica oscillatoria non è una difficoltà tecnica accessoria: è il cuore filosofico del problema, e merita di essere affrontata con la precisione che esige. Per farlo, occorre introdurre il concetto di **attrattore** nella sua accezione matematica rigorosa, e poi chiedersi se — e in quale senso — la grammatica possa essere concepita come un sistema di attrattori dello spazio degli stati corticali.

Dato un sistema dinamico autonomo descritto dall'equazione differenziale:

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n$$

un insieme $A \subset \mathbb{R}^n$ è un **attrattore** se soddisfa tre condizioni: è invariante sotto il flusso (se $x(0) \in A$ allora $x(t) \in A$ per ogni $t \geq 0$); possiede un bacino di attrazione aperto $B(A)$ tale che per ogni $x(0) \in B(A)$ si ha $d(x(t), A) \rightarrow 0$ per $t \rightarrow \infty$; ed è minimale, ossia non contiene sottoinsieme proprio che soddisfi le due condizioni precedenti. Gli attrattori più semplici sono i **punti fissi** (soluzioni stazionarie), i **cicli limite** (orbite periodiche isolate) e i **tori** (quasiperiodici). Ma un sistema non lineare può possedere attrattori geometricamente complessi, con dimensione frattale non intera: i cosiddetti **attrattori strani**, caratteristici del regime caotico, nei quali traiettorie vicine divergono esponenzialmente ma restano confinate in una regione limitata dello spazio degli stati.

Le oscillazioni neurali corrispondono tipicamente a cicli limite: l'attività ritmica di una popolazione è una traiettoria chiusa nello spazio degli stati, verso cui il sistema converge indipendentemente dalle condizioni iniziali (nell'intorno del bacino). L'accoppiamento tra oscillatori — formalmente descritto dal modello di Kuramoto esteso — corrisponde alla deformazione dello spazio degli attrattori: il paesaggio energetico del sistema accoppiato è diverso dalla semplice somma dei paesaggi dei sottosistemi isolati, e possono emergere nuovi attrattori — stati di sincronizzazione tra oscillatori a frequenze diverse — che non esistevano nel sistema disaccoppiato.

La domanda cruciale è: questo paesaggio di attrattori, che l'esperienza e la maturazione neurobiologica plasmano nel corso della vita del parlante, è strutturato in modo da rispecchiare le categorie sintattiche della grammatica della lingua acquisita? Ovvero: vi sono attrattori distinti — cicli limite, o tori, o strutture più complesse — corrispondenti alla differenza tra un sintagma verbale transitivo e uno intransitivo, tra una frase relativa soggetto e una relativa oggetto, tra una costruzione attiva e una passiva?

La risposta empiricamente plausibile è che la risposta sia affermativa, almeno in senso approssimativo. Studi di neurofisiologia computazionale — in particolare il lavoro di Fransen e collaboratori sui circuiti prefrontali per la memoria di lavoro, e il lavoro di Lundqvist e colleghi sulla codifica attractor-based delle rappresentazioni in lavoro — mostrano che reti corticali con connettività

appropriata possono mantenere rappresentazioni discrete stabili attraverso il meccanismo degli attrattori di punto fisso o di ciclo limite. La discretezza categoriale emerge come proprietà degli attrattori, non come proprietà delle variabili di stato elementari (potenziali di membrana, tassi di scarica): nessun singolo neurone è "il" neurone del sintagma nominale, ma la *configurazione* dell'intera rete può essere discretamente diversa tra la rappresentazione di un sintagma nominale e quella di un sintagma verbale.

Questo è precisamente il tipo di risposta che il funzionalismo multipla-realizzazione richiedeva: la categoria sintattica è implementata da una configurazione di rete, non da un elemento fisico specifico. Ma la novità che la prospettiva oscillatoria aggiunge è che la configurazione rilevante non è spaziale — quale neurone è attivo — bensì *temporale*: quale oscilla in fase con quale, a quale frequenza, con quale pattern di accoppiamento cross-frequenza.

XI. Il paesaggio energetico della grammatica: la funzione di Lyapunov

La caratterizzazione più rigorosa del paesaggio di attrattori di un sistema dinamico è quella fornita dalla **funzione di Lyapunov**, quando essa esiste. Per un sistema dissipativo con una funzione di Lyapunov $V: \mathbb{R}^n \rightarrow \mathbb{R}$, la condizione necessaria è:

$$dV/dt = \nabla V \cdot f(x) \leq 0 \quad \text{per ogni } x \in \mathbb{R}^n$$

con uguaglianza soltanto nei punti fissi. In questo caso, V è un potenziale decrescente lungo le traiettorie del sistema: il sistema "scende" nel paesaggio energetico definito da V fino a raggiungere un minimo locale — un attrattore di punto fisso.

Per sistemi corticali altamente non lineari e stocastici, la funzione di Lyapunov non esiste in generale in forma analitica chiusa, ma può essere approssimata attraverso diversi metodi. Il più elegante, nel contesto delle neuroscienze computazionali, è l'approccio della **energia libera variazionale** sviluppato da Karl Friston e collaboratori — che sarà discusso in dettaglio nella sezione successiva — nel quale la funzione di energia libera F gioca il ruolo di potenziale il cui gradiente discendente guida l'aggiornamento delle rappresentazioni interne del sistema.

Per il presente argomento, il punto essenziale è il seguente: se la grammatica corrisponde a un paesaggio di attrattori dello spazio degli stati corticale, allora l'acquisizione linguistica — il processo mediante il quale il bambino costruisce la propria grammatica — può essere concepita come il *plasmamento* progressivo di questo paesaggio attraverso l'esperienza linguistica. I bacini di attrazione delle categorie grammaticali si approfondiscono e si separano con l'esposizione

ripetuta a pattern linguistici della lingua target; le strutture non grammaticali corrispondono a regioni dello spazio degli stati con curvatura del potenziale positiva — colline, non valli — da cui il sistema viene repinto verso le strutture grammaticalmente lecite.

Questa immagine è concettualmente affascinante, ma è anche falsificabile in via di principio. Nei pazienti con disfasie acquisite — lesioni corticali che alterano selettivamente l'elaborazione sintattica — ci si dovrebbe aspettare un'alterazione del paesaggio energetico grammaticale: non l'abolizione di tutti gli attrattori, ma la fusione, l'appiattimento o la destabilizzazione di alcuni di essi. In termini oscillatori, questo si manifesterebbe come una riduzione del Modulation Index e del Phase-Locking Value nelle bande e nelle regioni corticali responsabili dell'elaborazione del tipo di struttura sintattica selettivamente compromessa.

Evidenze in questa direzione esistono, benché ancora frammentarie. Studi di EEG su pazienti afasici mostrano riduzioni del PAC nelle bande theta-gamma durante l'elaborazione di frasi con dipendenze a lungo raggio — proprio il tipo di struttura più frequentemente compromessa nelle afasie non fluenti. Questo risultato è compatibile con la tesi che il paesaggio oscillatorio non sia uniforme e che le lesioni corticali lo distorcano in modo strutturalmente specifico, non casuale.

XII. Invarianza e trasformazione: la grammatica come gruppo di simmetrie

Un modo alternativo e complementare di concepire la grammatica come struttura del sistema dinamico è attraverso il concetto di **simmetria** e il suo gruppo. Nelle teorie fisiche fondamentali — dalla meccanica classica alla meccanica quantistica, dalla relatività speciale alla teoria quantistica dei campi — le leggi di conservazione emergono da simmetrie del sistema: il teorema di Noether stabilisce una corrispondenza biunivoca tra ogni simmetria continua del funzionale d'azione e una grandezza conservata. Conservazione dell'energia ↔ invarianza per traslazione temporale; conservazione del momento ↔ invarianza per traslazione spaziale; conservazione del momento angolare ↔ invarianza per rotazione.

Il linguaggio, in quanto sistema formale, possiede anch'esso strutture di simmetria — benché di natura discreta, non continua, e dunque non soggette al teorema di Noether in senso stretto. Le trasformazioni sintattiche — passivizzazione, relativizzazione, formazione di domande, dislocazione a sinistra — sono operazioni che preservano certe relazioni strutturali mentre ne alterano altre. La passivizzazione, per esempio, preserva le relazioni tematiche (l'agente rimane agente) ma altera le relazioni grammaticali (il paziente diventa soggetto). La relativizzazione preserva la struttura argomentale del verbo della relativa ma estrae un argomento fuori dalla sua posizione originaria.

Nel Programma Minimalista, queste operazioni sono ricondotte a due meccanismi fondamentali: Merge esterno (che introduce nuovi elementi) e Merge interno (detto anche *Move*, che sposta un elemento da una posizione a un'altra lasciando una *traccia* o copia nella posizione di partenza). Merge interno è il meccanismo alla base di tutte le operazioni di movimento sintattico — e dunque, indirettamente, di tutta la fenomenologia delle dipendenze a lungo raggio che è uno dei tratti distintivi della sintassi umana rispetto alle grammatiche degli altri primati.

Dal punto di vista del sistema dinamico, Merge interno corrisponde a una **connessione a lungo raggio** tra due stati del sistema separati nel tempo: la posizione di origine dell'elemento mosso e la posizione in cui esso approda. Questa connessione temporale a lungo raggio richiede che l'informazione sulla posizione di partenza venga mantenuta attiva mentre il sistema processa il materiale intermedio — esattamente il ruolo funzionale che la memoria di lavoro svolge nell'elaborazione delle dipendenze a lungo raggio. E la firma oscillatoria della memoria di lavoro, come si è detto, è l'oscillazione beta: un'oscillazione di mantenimento che conserva attiva una rappresentazione attraverso i cicli successivi di elaborazione rapida (gamma).

Si profila così una corrispondenza suggestiva, che merita di essere enunciata con precisione: le dipendenze sintattiche a lungo raggio — prodotto di Merge interno — si rispecchiano nelle dipendenze a lungo raggio *temporale* dell'oscillazione beta, che mantiene la traccia dell'elemento estratto attraverso i cicli theta e gamma successivi. La struttura delle simmetrie sintattiche — quali elementi possono essere estratti, da dove, verso dove — si riflette nella struttura delle connessioni a lungo raggio tra oscillatori neurali che l'esperienza linguistica ha selezionato e consolidato.

Questa corrispondenza non è ancora una teoria compiuta: mancano i passaggi tecnici che la rendano quantitativamente precisa e falsificabile. Ma fornisce un'intuizione profonda su dove cercare: non nelle mappe topografiche della corteccia, non nelle singole regioni di Broca e Wernicke, ma nella topologia dello spazio degli attrattori e nella geometria delle connessioni tra oscillatori che queste aree mediano.

XIII. La competenza come invariante del processo: oltre Chomsky

La tradizione generativista, da Chomsky in poi, ha sempre insistito sulla necessità di distinguere la competenza — la conoscenza implicita della grammatica, astratta e ideale — dall'esecuzione — il processamento linguistico reale, soggetto a limiti di memoria, distrazioni, errori, variazioni individuali. Questa distinzione aveva una precisa funzione metodologica: isolare l'oggetto di studio della linguistica (la grammatica come sistema formale) dalla psicologia del

processamento e dalle variabili situazionali che complicano l'osservazione empirica.

La proposta oscillatoria, nella sua versione più radicale, non abolisce questa distinzione ma la reinterpreta in termini che la rendono empiricamente agganciata. La competenza grammaticale non è un sistema separato che si applica, come un software esterno, al processamento: è l'**invariante strutturale** del sistema dinamico, ossia la proprietà del sistema che rimane stabile attraverso le istanze diverse di elaborazione della frase.

Formalmente, si può concepire la competenza come l'insieme degli attrattori del sistema dinamico corticale e la struttura delle loro relazioni: quali stati sono attrattori (strutture grammaticali lecite), quali sono regioni di repulsione (strutture illecite), quali attrattori sono connessi da traiettorie accessibili (derivazioni grammaticali valide) e quali non lo sono. L'esecuzione, viceversa, corrisponde alle traiettorie specifiche che il sistema percorre nello spazio degli stati durante l'elaborazione di frasi particolari — traiettorie che dipendono dalle condizioni iniziali, dal rumore stocastico, dalla configurazione della memoria di lavoro, dalla complessità locale della struttura incontrata.

In questa prospettiva, la distinzione tra competenza ed esecuzione sopravvive, ma cessa di essere una distinzione tra un livello astratto-formale e un livello concreto-fisico. Diventa invece una distinzione tra due aspetti dello stesso sistema dinamico: la struttura del paesaggio energetico (competenza) e la dinamica delle traiettorie che quel paesaggio determina (esecuzione). Questa reinterpretazione ha conseguenze non banali.

La prima è che la competenza grammaticale non è necessariamente consapevole, né necessariamente accessibile all'introspezione, né necessariamente modulare nel senso di Fodor: è semplicemente la struttura degli attrattori di un sistema dinamico distribuito, e il suo essere "conoscenza" è una proprietà funzionale, non una proprietà intrinseca. Il parlante "sa" la grammatica nel senso in cui un fluido "sa" la termodinamica: non perché possieda una rappresentazione esplicita delle sue leggi, ma perché il suo comportamento è sistematicamente conforme ad esse, governato com'è da una struttura interna che quelle leggi descrivono.

La seconda conseguenza riguarda la variazione linguistica individuale. Se la competenza è la struttura degli attrattori del sistema dinamico, allora differenze individuali nella sintassi — nell'accettabilità di certi costrutti, nella sensibilità a certi tipi di ambiguità, nelle soglie di accettazione di dipendenze a lungo raggio — corrispondono a differenze nella geometria del paesaggio energetico. Queste differenze potrebbero essere quantificate, in linea di principio, attraverso misure di PAC e PLV che caratterizzano l'architettura oscillatoria individuale.

La terza e più sottile conseguenza riguarda il rapporto tra grammatica e significato. Nella concezione minimalista standard, la sintassi opera su oggetti formali a cui il lessico assegna contenuto semantico alle interfacce: la struttura

sintattica è computazionalmente cieca al significato, anche se genera le strutture che permettono l'interpretazione semantica. Ma se la competenza sintattica è la struttura degli attrattori del sistema dinamico, e se le stesse reti oscillatorie che realizzano la gerarchia sintattica sono connesse — attraverso il beta e il gamma — con le reti semantiche e concettuali, allora la separazione netta tra sintassi e semantica diventa empiricamente molto più difficile da sostenere. Non perché la sintassi non esista come sistema relativamente autonomo, ma perché i suoi attrattori potrebbero essere modellati, fin dall'acquisizione, dall'esperienza semantica e dall'uso comunicativo — non essere dunque strutture puramente formali preformate, ma proprietà emergenti da un sistema che ha imparato a sintetizzare struttura e significato in modo indissolubile.

Questo è il punto in cui la proposta oscillatoria converge naturalmente con le teorie della codifica predittiva, che affrontano il problema dell'integrazione semantico-sintattica da una prospettiva complementare e ugualmente formalmente potente.

XIV. Il cervello come macchina di previsione: fondamenti della codifica predittiva

Il paradigma della **codifica predittiva** — detto anche *predictive processing* nella versione più ampia elaborata da Andy Clark e Jakob Hohwy — costituisce uno degli sviluppi teorici più fecondi delle neuroscienze computazionali degli ultimi due decenni. Nella sua formulazione più generale, esso propone che il cervello non sia fondamentalmente un organo di rilevazione passiva degli stimoli, ma una macchina generativa che produce incessantemente predizioni sul proprio input sensoriale e aggiorna queste predizioni sulla base degli errori di previsione. L'elaborazione percettiva, motoria e cognitiva viene ricondotta a un unico principio computazionale: la minimizzazione dell'errore di previsione, o, nella formulazione più astratta di Friston, la minimizzazione della **energia libera variazionale**.

Occorre qui presentare il formalismo con la precisione necessaria, poiché la sua potenza esplicativa dipende interamente dalla sua struttura matematica, e una presentazione vaga ne tradirebbe sia lo spirito che le implicazioni.

Si denota con $s \in S$ il vettore degli stati del mondo (nascosti e non direttamente osservabili) e con $o \in O$ le osservazioni (l'input sensoriale). Il cervello possiede un **modello generativo** del mondo, espresso come una distribuzione di probabilità congiunta $p(o, s) = p(o | s) \cdot p(s)$, dove $p(o | s)$ è la verosimiglianza (come il mondo produce le osservazioni dato lo stato) e $p(s)$ è il prior (le credenze a priori sullo stato del mondo, indipendenti dall'osservazione corrente).

Il problema dell'inferenza percettiva è quello del calcolo della distribuzione *a posteriori* $p(s | o) \propto p(o | s) \cdot p(s)$, ossia la credenza sullo stato del mondo

aggiornata all'osservazione corrente. Per sistemi complessi, questo calcolo è generalmente intrattabile: richiederebbe l'integrazione su tutti i possibili stati del mondo, uno spazio di dimensionalità altissima.

L'approssimazione variazionale sostituisce l'inferenza esatta con la ricerca della distribuzione $q(s)$ — in una classe di distribuzioni trattabili, ad esempio gaussiane — che minimizza la divergenza di Kullback-Leibler da $p(s | o)$:

$$q^*(s) = \operatorname{argmin}_{\{q\}} D_{\{KL\}}(q(s) \parallel p(s | o))$$

Attraverso la decomposizione della divergenza KL, si mostra che minimizzare $D_{\{KL\}}(q \parallel p(\cdot | o))$ equivale a minimizzare la quantità:

$$\begin{aligned} F[q] &= \mathbb{E}_q[\log q(s)] - \mathbb{E}_q[\log p(o, s)] \\ &= D_{\{KL\}}(q(s) \parallel p(s)) - \mathbb{E}_q[\log p(o | s)] \end{aligned}$$

Il primo termine è la divergenza KL tra la distribuzione posteriore approssimata $q(s)$ e il prior $p(s)$: una misura della "sorpresa" relativa rispetto alle credenze precedenti. Il secondo termine, cambiato di segno, è la verosimiglianza attesa: quanto bene il modello spiega le osservazioni sotto la distribuzione q . La somma $F[q]$ è l'**energia libera variazionale**, il cui minimo fornisce la miglior approssimazione trattabile dell'inferenza bayesiana esatta.

Una proprietà cruciale di F è che $F[q] \geq -\log p(o)$ per ogni q , con uguaglianza quando $q(s) = p(s | o)$. Dunque l'energia libera è un **upper bound** sulla sorpresa (o auto-informazione) $-\log p(o)$: minimizzare F corrisponde a minimizzare la sorpresa attesa sulle osservazioni. Il cervello come macchina predittiva è dunque un sistema che minimizza la propria sorpresa circa il mondo — e lo fa aggiornando le sue rappresentazioni interne (minimizzazione dell'energia libera percettiva) oppure agendo sul mondo per renderlo più conforme alle proprie predizioni (minimizzazione dell'energia libera attiva, che fonda la teoria del controllo motorio e dell'azione).

XV. La gerarchia predittiva e il suo substrato oscillatorio

Ciò che rende il paradigma della codifica predittiva particolarmente rilevante per il problema della sintassi è la sua struttura **gerarchica**. Nella formulazione di Friston — e nei modelli di elaborazione visiva di Rao e Ballard che l'hanno preceduta — il sistema nervoso è concepito come una gerarchia di livelli, ciascuno dei quali genera predizioni sul livello immediatamente inferiore e riceve segnali di errore di previsione dal medesimo.

Formalmente, si considera una gerarchia a L livelli, con stati $x^{(l)}$ al livello l , per $l = 1, \dots, L$. Ogni livello genera una previsione del livello sottostante attraverso una

funzione generativa $g^{(l)}: x^{(l+1)} \rightarrow x^{(l)}$, e riceve un segnale di errore $\epsilon^{(l)} = x^{(l)} - g^{(l)}(x^{(l+1)})$. Le equazioni di aggiornamento della dinamica predittiva sono:

$$\dot{x}^{(l)} = -\epsilon^{(l)} + (\partial g^{(l)} / \partial x^{(l)})^T \cdot \epsilon^{(l-1)} \quad \text{per } l = 2, \dots, L-1$$

Queste equazioni catturano la struttura bidirezionale delle connessioni corticali: i segnali discendenti (*top-down*) trasportano le predizioni generate dai livelli superiori; i segnali ascendenti (*bottom-up*) trasportano gli errori di previsione, le deviazioni tra predizioni e input effettivi. In questa architettura, l'elaborazione sensoriale non è un flusso unidirezionale dall'input alla risposta: è un processo iterativo di riconciliazione tra predizioni generate dall'alto e segnali di errore risalenti dal basso.

La connessione con le oscillazioni corticali è il punto in cui le due grandi teorie computazionali qui considerate si incontrano in modo non banale. Un'ipotesi influente — avanzata da Friston stesso e indipendentemente da Bastos e collaboratori in un articolo del 2012 su *Neuron* — è che le connessioni discendenti (predizioni) siano mediate prevalentemente dalle oscillazioni **alpha-beta** (8–30 Hz), mentre le connessioni ascendenti (errori di previsione) siano mediate prevalentemente dalle oscillazioni **gamma** (30–80 Hz). Questa dissociazione funzionale corrisponde alla dissociazione anatomica tra le connessioni discendenti che originano prevalentemente dagli strati superficiali della corteccia e le connessioni ascendenti che originano dagli strati granulari (strato IV), con proprietà biofisiche differenti che favoriscono frequenze oscillatorie distinte.

Se questa ipotesi è corretta — e l'evidenza sperimentale si accumula nella sua direzione, benché non sia ancora definitiva — allora l'architettura oscillatoria corticale è anche l'architettura computazionale della gerarchia predittiva: le bande di frequenza non sono soltanto scale temporali di integrazione, ma anche direzionalità di flusso informativo. Il beta trasporta le ipotesi dall'alto verso il basso; il gamma riporta le deviazioni dal basso verso l'alto. Il delta-theta, nell'estensione di questa proposta al dominio linguistico, organizza temporalmente l'intera gerarchia su scale che corrispondono ai livelli strutturali della frase.

XVI. Il grounding semantico nella gerarchia oscillatoria predittiva

L'integrazione della codifica predittiva con la proposta oscillatoria permette di affrontare un problema che la sola prospettiva oscillatoria lasciava irrisolto: il **grounding semantico**, ossia il modo in cui le strutture sintattiche si connettono con il contenuto concettuale e referenziale del linguaggio.

Nella tradizione formale, il significato delle frasi è compositivo: il significato dell'intero si ottiene combinando i significati dei costituenti secondo regole determinate dalla struttura sintattica. Il principio di composizionalità — attribuito a Frege, benché la sua formulazione precisa sia posteriore e controversa — stabilisce che la funzione semantica della sintassi è di proiettare, attraverso le proprie operazioni, la struttura degli argomenti e dei predicati. La forma logica, in senso generativista, è precisamente la struttura sintattica che l'interfaccia LF espone all'interpretazione semantica.

Nella prospettiva della codifica predittiva, il significato non è una proprietà assegnata alle strutture dall'esterno mediante regole di interpretazione: è parte integrante del processo generativo mediante il quale il cervello costruisce il proprio modello del discorso in corso. Ogni parola incontrata aggiorna la distribuzione di probabilità sugli stati del mondo (e del discorso) che il modello generativo mantiene; la struttura sintattica vincola e dirige questo aggiornamento, determinando quali stati sono compatibili con l'input corrente e quali devono essere soppressi.

In termini formali, si può concepire il modello generativo del linguaggio come una distribuzione di probabilità su tuple (struttura sintattica, significato, contesto):

$$p(o, s, m, c) = p(o \mid s) \cdot p(s \mid m, c) \cdot p(m \mid c) \cdot p(c)$$

dove o è il segnale acustico osservato, s è la struttura sintattica, m è il contenuto semantico, e c è il contesto discorsivo. La struttura sintattica non è semplicemente il tramite meccanico attraverso cui scorrono le parole verso l'interpretazione: è una variabile latente del modello generativo, inferita congiuntamente al contenuto semantico e al contesto, vincolando e vincolata da entrambi.

Il **prior sintattico** $p(s \mid m, c)$ cattura la probabilità a priori di una struttura sintattica dato un contenuto semantico e un contesto: certe strutture sono più probabili per certi tipi di eventi (le frasi causali sono più probabili nei contesti di spiegazione; le domande polari nei contesti di negoziazione dell'informazione). Il **prior semantico** $p(m \mid c)$ cattura le aspettative concettuali sul contenuto del discorso dato il contesto.

L'accoppiamento tra questi due livelli di prior — sintattico e semantico — è ciò che il paradigma della codifica predittiva illumina in modo particolarmente perspicuo. La celebre **N400**, la componente negativa dell'evento-correlato cerebrale (*Event-Related Potential*, ERP) che si osserva circa 400 millisecondi dopo l'incontro con una parola semanticamente anomala nel suo contesto, è interpretata nella prospettiva predittiva come il segnale di errore di previsione semantica — la scarica verso l'alto di $e^{\wedge}(m)$, l'errore al livello semantico della gerarchia. La **P600**, che compare circa 600 millisecondi dopo una violazione

sintattica, corrisponde all'errore di previsione al livello sintattico — $\epsilon^s(s)$. La dissociazione temporale tra i due segnali suggerisce che i due livelli operino con dinamiche temporali diverse, compatibili con la loro realizzazione a bande di frequenza distinte nella gerarchia oscillatoria.

Ma la relazione tra N400 e P600 è più sottile di quanto questa dissociazione suggerisca. In condizioni di anomalia semantica in contesto sintattico regolare — *"Il professore beve il sonetto"* — si osserva una N400 elevata. In condizioni di anomalia sintattica in contesto semantico coerente — un verbo coniugato in modo erraneo in una frase altrimenti sensata — si osserva una P600 elevata. Ma in condizioni in cui la frase è sia semanticamente che sintatticamente anomala, o in cui l'anomalia sintattica ha conseguenze semantiche, i due segnali si sovrappongono e interagiscono in modo non lineare, suggerendo che i due livelli della gerarchia non siano indipendenti ma si co-aggiornino in modo reciproco.

Questa non-indipendenza è perfettamente coerente con la struttura della gerarchia predittiva: i livelli superiori aggiornano i livelli inferiori attraverso le connessioni discendenti (modulate dal beta), e i livelli inferiori aggiornano i livelli superiori attraverso i segnali di errore ascendenti (modulati dal gamma). L'integrazione semantico-sintattica non è sequenziale — prima la sintassi, poi il semantico, o viceversa — ma parallela e iterativa, convergente verso lo stato di equilibrio del sistema gerarchico che minimizza congiuntamente gli errori di previsione a tutti i livelli.

XVII. Il problema della referenza e l'identità attraverso il discorso

Vi è un ulteriore aspetto della semantica che la prospettiva puramente oscillatoria fatica a cogliere, ma che la codifica predittiva illumina in modo inatteso: il problema della **referenza** e dell'identità degli referenti attraverso il discorso. Quando si dice "Mario entrò nella stanza. Lui era esausto", il pronome "lui" è interpretato come coreferente con "Mario", benché tra di loro intercorra un'intera frase. Questa identità di riferimento non è determinata dalla struttura sintattica locale — nella seconda frase, "lui" è semplicemente un sintagma pronominale in posizione soggetto — ma da meccanismi di anafora che operano a livello del modello del discorso: la struttura di entità, proprietà e relazioni che il parlante-ascoltatore mantiene aggiornata attraverso la conversazione.

Nel framework della codifica predittiva, il modello del discorso è parte del modello generativo ai livelli più alti della gerarchia: è la struttura di prior $p(c)$ che determina le probabilità di occorrenza di certi referenti, certi eventi, certe relazioni. La risoluzione dell'anafora corrisponde all'aggiornamento di questo prior alla luce dell'input corrente: tra tutti i referenti nel modello del discorso, quale ha la probabilità *a posteriori* più elevata di essere l'antecedente del pronome, date le sue proprietà di accordo, la sua posizione nell'albero del discorso, e le aspettative tematiche sul prosieguo?

In termini oscillatori, si può ipotizzare che il mantenimento del modello del discorso — la struttura di entità e relazioni che si estende attraverso molte frasi — sia supportato da oscillazioni a frequenza molto bassa, nell'intervallo infra-delta (< 0.5 Hz) o persino nei ritmi ultralenti dei processi corticali (*slow cortical potentials*). Il modello del discorso è, in questo senso, il livello più alto della gerarchia oscillatoria: l'impalcatura temporale più ampia entro cui si inscrivono le gerarchie sintattiche di singole frasi. Le frasi sono ai periodi discorsivi ciò che le sillabe sono alle frasi: unità annidate dentro unità più grandi, ciascuna realizzata da un livello della gerarchia oscillatoria corticale.

Questa estensione — dalla frase al discorso, dalla gerarchia sintattica alla struttura discorsiva — non è ancora oggetto di ricerca sistematica nel dominio oscillatorio. È, in senso proprio, un programma di ricerca aperto: e questa apertura è forse il maggiore indizio della sua fecondità. Un'idea che esaurisce i propri problemi nell'atto stesso di essere enunciata non merita il nome di teoria.

XVIII. Linguaggio, oscillazione e la domanda sulla natura del senso

Si giunge così a un punto in cui le traiettorie fin qui percorse — la teoria sintattica formale, la neurofisiologia delle oscillazioni, la codifica predittiva, la fenomenologia husserliana — convergono verso una domanda che nessuna di esse, presa isolatamente, è in grado di formulare con piena chiarezza: che cosa significa *comprendere* una frase?

La comprensione non è riducibile né alla corretta costruzione della struttura sintattica, né all'aggiornamento del modello semantico, né alla minimizzazione dell'energia libera. È qualcosa che, per quanto l'analisi proceda, continua a eccedere le descrizioni parziali che di essa si possono dare. Wittgenstein, nelle *Ricerche filosofiche*, scriveva che nel comprendere una parola avviene qualcosa di caratteristico — un'esperienza di riconoscimento immediato, un Aha-Erlebnis, il sentimento che le cose vadano a posto — ma che questo qualcosa è irriducibile alla somma delle operazioni che lo rendono possibile. Il significato non è nel testo, non è nel lettore, non è nella struttura: è nell'*uso*, nella pratica sociale condivisa che rende la struttura applicabile e la comprensione verificabile.

Questa insistenza wittgensteiniana sull'uso non è una rinuncia all'analisi scientifica: è una messa in guardia contro la tentazione di identificare il livello di spiegazione con il fenomeno da spiegare. La gerarchia oscillatoria, per quanto precisamente modellata e sperimentalmente confermata, descrive le condizioni neurofisiologiche che rendono possibile la comprensione linguistica — non la comprensione stessa, nella sua irriducibile dimensione di senso condiviso. Tra il modello dinamico del sistema corticale e l'esperienza di capire una frase vi è uno iato che nessuna riduzione, per quanto sofisticata, riesce a colmare — non perché la riduzione sia impossibile in linea di principio, ma perché i livelli di descrizione sono eterogenei e non intertraducibili senza residuo.

Questo residuo non è un errore da eliminare: è la traccia dell'irriducibilità del senso alla struttura, del significato alla forma, dell'esperienza alla funzione. È, in definitiva, ciò che rende il linguaggio un oggetto inesauribile per la filosofia — non nonostante la crescente sofisticazione delle spiegazioni scientifiche che ne rendiamo, ma proprio grazie ad essa: ogni livello di spiegazione che si aggiunge fa emergere con più nitidezza il profilo di ciò che resta inspiegato.

Il cervello che oscilla, che si sincronizza, che minimizza la propria sorpresa circa il mondo, che mantiene il modello del discorso nell'impalcatura delle sue oscillazioni più lente — questo cervello è uno degli oggetti più straordinari che l'universo abbia prodotto. Ma il parlante che comprende, che ride di una battuta, che si commuove a una metafora, che intuisce l'ironia in un tono di voce: questo parlante è qualcosa che il cervello *fa*, non qualcosa che il cervello *è*. E questa differenza — tra il fare e l'essere, tra il processo e il suo senso — resterà, con ogni probabilità, il problema irrisolto al cuore di ogni futura neurosociologia del linguaggio, per quanto elaborata e matematicamente potente essa possa diventare.

XIX. Il problema della variazione: lingue diverse, cervelli diversi?

Una delle obiezioni più immediate che si può rivolgere alla proposta oscillatoria è la seguente: se la gerarchia sintattica è realizzata dall'architettura di accoppiamento cross-frequenza, e se questa architettura è plasmata dall'esperienza linguistica, allora parlanti di lingue tipologicamente diverse dovrebbero esibire architetture oscillatorie diverse. E poiché le lingue del mondo variano enormemente nei loro parametri strutturali — nell'ordine delle parole, nel tipo di morfologia, nella presenza o assenza di classi tonali, nella struttura prosodica — ci si aspetterebbe una corrispondente variazione nelle firme oscillatorie dell'elaborazione linguistica. Ma quanto questa variazione è documentata, e quanto è compatibile con la tesi di una gerarchia oscillatoria universale che fondi la facoltà linguistica nel senso biolinguistico?

La risposta richiede di distinguere con cura tra due livelli della proposta oscillatoria: il livello degli **universali** — la gerarchia delta-theta-beta-gamma come architettura neurale della facoltà linguistica in senso lato — e il livello dei **parametri** — le specifiche modalità di accoppiamento, i range precisi di frequenza, le asimmetrie emisferiche, che variano in funzione della struttura della lingua acquisita. Il parallelismo con la teoria dei Principi e Parametri di Chomsky è deliberato: così come il minimalismo distingue tra principi computazionali universali (Merge, il principio di dipendenza dalla struttura) e parametri di variazione che le lingue impostano diversamente (il valore del parametro pro-drop, l'ordine di testa e complemento, la morfologia del caso), la proposta oscillatoria distingue tra l'architettura gerarchica del coupling come universale biologico e i dettagli del suo settaggio come prodotto dell'esperienza linguistica specifica.

L'universale più robusto che emerge dagli studi cross-linguistici è la **sincronia sillabica theta**: in tutte le lingue esaminate — lingue romanze, germaniche, semitiche, cinesi, giapponese — le oscillazioni corticali nella banda theta mostrano una modulazione sistematica al ritmo sillabico del parlato. Questo risultato è stato replicato in decine di laboratori e su popolazioni diverse, ed è oggi considerato uno dei fatti più solidi della neurolinguistica oscillatoria. La sillaba sembra essere l'unità universale di processing al confine tra la fonetica e la fonologia, e la sua realizzazione oscillatoria è parimenti universale: non perché le sillabe siano uguali in tutte le lingue — variano enormemente in struttura, durata, complessità consonantica — ma perché l'elaborazione che il cervello compie su di esse si svolge sempre nella stessa finestra temporale theta.

XX. Lingue tonali e la quarta dimensione prosodica

Le **lingue tonali** — un gruppo che comprende circa il 70% delle lingue del mondo, tra cui le lingue della famiglia sino-tibetana, le lingue bantu, molte lingue amerindiane e austronesiane — pongono una sfida particolarmente interessante alla proposta oscillatoria. In queste lingue, il tono — la frequenza fondamentale (F0) del suono vocalico — è un tratto distintivo lessicale: la stessa sequenza di segmenti può avere significati completamente diversi a seconda del profilo tonale. Nel mandarino standard, la sillaba /ma/ con tono alto e piatto significa "madre", con tono ascendente significa "canapa", con tono basso e cadente-ascendente significa "cavallo", con tono cadente significa "sgridare". Il tono non è dunque, in queste lingue, una proprietà prosodica sovrapposta su una struttura lessicale altrimenti determinata: è parte integrante della struttura fonologica del lessema.

Il segnale acustico del tono — la curva di F0 — ha una struttura temporale che si dispiega tipicamente su scale di 150–300 millisecondi, sovrapponendosi alla scala theta dell'elaborazione sillabica. Si pone dunque immediatamente la domanda: il cervello del parlante di lingua tonale usa le stesse oscillazioni theta per elaborare sia la struttura sillabica che il profilo tonale? O le due elaborazioni sono dissociate in bande di frequenza diverse?

La ricerca disponibile — in particolare gli studi di Ding e collaboratori su parlanti madrelingua mandarino — suggerisce una soluzione elegante: i parlanti di lingue tonali sviluppano una specializzazione oscillatoria nell'elaborazione della curva di F0 che si sovrappone parzialmente ma non si identifica con l'elaborazione sillabica. Le oscillazioni theta continuano a sincronizzarsi con il ritmo sillabico, ma le oscillazioni nella banda **alpha bassa** (circa 8–10 Hz) mostrano una modulazione aggiuntiva correlata con il profilo tonale. Il tono lessicale, in altre parole, recluta una banda di frequenza supplementare — o, più precisamente, una modulazione aggiuntiva di una banda già esistente — che si aggiunge alla gerarchia oscillatoria standard senza alterarla nella sua architettura fondamentale.

Formalmente, si può descrivere questa situazione come un ampliamento del modello di accoppiamento: per le lingue tonali, l'operatore di accoppiamento cross-frequenza include un termine aggiuntivo che lega le oscillazioni theta (sillabiche) alle oscillazioni alpha (tonali) attraverso un meccanismo di accoppiamento fase-fase (*phase-phase coupling*), distinto dal coupling fase-ampiezza che organizza la gerarchia sillaba-fonema. L'equazione di fase per l'oscillatore tonale $\theta_{\text{tone}}(t)$ assume la forma:

$$d\theta_{\text{tone}}/dt = \omega_{\text{tone}} + K_{\text{tone-syl}} \cdot \sin(n \cdot \theta_{\text{syl}} - \theta_{\text{tone}}) + \eta(t)$$

dove θ_{syl} è la fase dell'oscillatore sillabico (theta), n è il rapporto di frequenza intero tra il ritmo sillabico e il ritmo tonale, e $\eta(t)$ è il rumore stocastico. Quando $K_{\text{tone-syl}}$ è sufficientemente elevato, l'oscillatore tonale si sincronizza con l'armonico n -esimo dell'oscillatore sillabico, producendo uno stato di *n:1 phase locking*. Questo stato di sincronizzazione è precisamente quello che permette al sistema di integrare l'informazione tonale con quella sillabica in modo temporalmente coerente — e la sua insorgenza durante lo sviluppo coincide, negli studi longitudinali disponibili, con il periodo in cui i bambini che acquisiscono lingue tonali cominciano a trattare il tono come un tratto lessicale anziché come variazione prosodica.

La conseguenza teorica più importante di questo risultato è che la variazione tipologica nel dominio tonale non rifà il progetto dell'architettura oscillatoria: lo *estende*. La gerarchia delta-theta-beta-gamma rimane l'impalcatura fondamentale; le lingue tonali aggiungono un ulteriore strato di accoppiamento che realizza la dimensione prosodico-lessicale specifica di queste lingue. L'universale è l'architettura della gerarchia; il parametro è la presenza o assenza (e il tipo) di questo strato aggiuntivo.

XXI. Lingue a ordine libero e la prosodia come sintassi

Un secondo banco di prova cruciale per la proposta oscillatoria è rappresentato dalle lingue a **ordine delle parole relativamente libero**: lingue come il latino classico, il russo, il giapponese, il turco, in cui la sequenza lineare degli elementi nella frase è relativamente non vincolata dalla struttura sintattica e può variare in funzione di fattori pragmatici — tema-rema, focus, topic — senza alterare le relazioni grammaticali di base.

In queste lingue, le relazioni grammaticali sono codificate morfologicamente (tipicamente attraverso il caso) piuttosto che posizionalmente. Il latino *Canem vidit Marcus* e *Marcus vidit canem* e *Vidit Marcus canem* esprimono la stessa proposizione — Marco vide il cane — attraverso ordini diversi, e la differenza tra le varianti è pragmatica, non semantica nel senso delle relazioni tematiche. Il

caso accusativo del nome *canem* segnala univocamente il ruolo di paziente, indipendentemente dalla posizione nella stringa.

Questo fatto pone un problema interessante per la proposta oscillatoria: se le dipendenze sintattiche sono realizzate da connessioni temporali a lungo raggio tra oscillatori, e se l'ordine lineare è relativamente arbitrario rispetto alla struttura sintattica, allora il sistema oscillatorio non può affidarsi alla posizione temporale degli elementi nella stringa per determinare le loro relazioni strutturali. Come vengono allora stabilite le connessioni oscillatorie tra testa e dipendente quando il dipendente può trovarsi prima o dopo la testa con uguale frequenza?

Una risposta plausibile — ancora largamente ipotetica in assenza di evidenza elettrofisiologica diretta su lingue a caso — è che nelle lingue morfologicamente ricche la **prosodia** svolga un ruolo strutturale più marcato che nelle lingue a ordine fisso. La prosodia — l'organizzazione del parlato in unità intonative, gruppi di respiro, unità di tono — è universalmente presente e universalmente realizzata da oscillazioni delta (per le frasi intonative) e theta (per le parole fonologiche). Nelle lingue a ordine libero, la struttura prosodica potrebbe fornire l'impalcatura oscillatoria che nelle lingue a ordine fisso è parzialmente fornita dall'ordine posizionale: i costituenti che formano un gruppo sintattico tendono a essere raggruppati prosodicamente, e questo raggruppamento è realizzato dall'accoppiamento oscillatorio delta-theta che marca i confini dei sintagmi intonativi.

Questa ipotesi è compatibile con un fatto tipologico ben noto: le lingue a ordine libero tendono ad avere una struttura prosodica più rigida e funzionalmente più carica rispetto alle lingue a ordine fisso. Il russo, ad esempio, ha vincoli prosodici molto più severi del polacco su certi tipi di focus e di topic; il turco usa sistematicamente la prosodia per marcare la struttura informativa che l'inglese marca con l'ordine delle parole. Se la prosodia è la struttura oscillatoria delta-theta, allora le lingue a ordine libero sarebbero caratterizzate da un uso più intenso di queste frequenze come strumento di codifica strutturale — un uso che le lingue a ordine fisso possono in parte delegare alla posizione temporale degli elementi nella stringa.

XXII. L'asimmetria emisferica e la lateralizzazione del linguaggio

Nessuna discussione neurofisiologica del linguaggio può permettersi di ignorare il fatto più robusto e replicato di tutta la neuropsicologia: la **lateralizzazione** dell'elaborazione linguistica all'emisfero sinistro. In circa il 96% dei destrimani e nel 70–80% dei mancini, le aree cerebrali critiche per la produzione e la comprensione del linguaggio sono prevalentemente localizzate nell'emisfero sinistro, in corrispondenza anatomica con l'area di Broca (giro frontale inferiore sinistro, BA 44/45) e l'area di Wernicke (giro temporale superiore posteriore sinistro, BA 22).

Questa asimmetria è tanto più misteriosa quanto più si considera che i due emisferi sono anatomicamente quasi simmetrici, connessi dal corpo calloso, e che entrambi sono capaci di elaborazione linguistica rudimentale. Perché una funzione cognitiva così complessa dovrebbe essere così drasticamente lateralizzata? E qual è il substrato di questa asimmetria?

Una risposta che la proposta oscillatoria rende disponibile in modo particolarmente elegante è la seguente: l'asimmetria emisferica nel linguaggio riflette un'asimmetria nelle proprietà temporali delle oscillazioni dei due emisferi. Studi di EEG e MEG hanno documentato che la corteccia temporale superiore sinistra e quella destra differiscono sistematicamente nel loro profilo di oscillazioni spontanee: l'emisfero sinistro mostra un'oscillazione theta più pronunciata e più strettamente sincronizzata con il ritmo sillabico del parlato, mentre l'emisfero destro mostra oscillazioni delta più pronunciate e più strettamente sincronizzate con il ritmo prosodico sovrasillabico — le frasi intonative, i gruppi di respiro.

Questa dissociazione ha una precisa funzione computazionale nella teoria di Poeppel: l'emisfero sinistro specializza l'elaborazione a grana fine (sillabe, morfemi, parole), mentre l'emisfero destro specializza l'elaborazione a grana grossa (frasi, clausole, struttura prosodica). I due emisferi operano in parallelo su scale temporali diverse e si integrano attraverso le connessioni callosali, producendo la ricca comprensione del linguaggio naturale che nessun emisfero da solo potrebbe realizzare.

Questa proposta, nota come ipotesi della **Asymmetric Sampling in Time** (AST), può essere formalizzata come una differenza nelle distribuzioni di frequenza naturale degli oscillatori nei due emisferi. Se si denota con $g_L(\omega)$ e $g_R(\omega)$ le distribuzioni di frequenza naturale degli oscillatori corticali nell'emisfero sinistro e destro rispettivamente, l'ipotesi AST predice:

$g_L(\omega)$ ha un picco prominente nell'intervallo theta (4–8 Hz)

$g_R(\omega)$ ha un picco prominente nell'intervallo delta (1–4 Hz)

con la conseguenza che, nel regime vicino alla transizione critica del modello di Kuramoto, la frequenza dominante della sincronizzazione collettiva è diversa nei due emisferi. L'emisfero sinistro entra in sincronia theta preferenzialmente; l'emisfero destro in sincronia delta preferenzialmente. La finestra temporale di integrazione — l'intervallo entro cui gli input vengono trattati come appartenenti allo stesso evento percettivo — è dunque più corta a sinistra (circa 25–250 millisecondi, la scala theta) e più lunga a destra (circa 250–2000 millisecondi, la scala delta).

La lateralizzazione del linguaggio non è dunque un mistero irriducibile ma una conseguenza prevedibile di questa asimmetria temporale: poiché l'analisi segmentale e morfolessicale richiede finestre di integrazione brevi (scala sillabica,

theta), è naturalmente delegata all'emisfero con la finestra più breve — il sinistro. L'analisi prosodica e discorsiva, che richiede finestre di integrazione lunghe (scala frasale, delta), è naturalmente delegata all'emisfero con la finestra più ampia — il destro.

Questa cornice spiega anche perché la lateralizzazione non è assoluta: entrambi gli emisferi partecipano al linguaggio, ma ciascuno contribuisce in modo privilegiato alla scala temporale per cui è meglio attrezzato. E spiega perché pazienti con lesioni dell'emisfero destro spesso conservano una comprensione grammaticale di base ma perdono la sensibilità alle sfumature prosodiche, all'ironia, alla metafora estesa su lunghi brani di testo — tutte funzioni che richiedono finestre di integrazione delta o infra-delta.

XXIII. Ritmo e metro: le lingue poetiche come esperimento naturale

Un dominio di variazione tipologica che la proposta oscillatoria illumina in modo inatteso è quello delle **tradizioni poetiche** e delle loro regole metriche. La poesia — in quasi tutte le culture che ne hanno sviluppato una forma — organizza il testo verbale secondo pattern ritmici regolari che impongono al linguaggio una struttura temporale più vincolata di quella della prosa. Il metro greco classico, la prosodia quantitativa latina, il verso accentuativo germanico, il metro sillabico delle tradizioni romanze, il metro sillabo-tonico del verso russo: ciascuna di queste tradizioni prescrive una struttura ritmica specifica, che il parlante-ascoltatore percepisce come soddisfatta o violata.

Dal punto di vista della proposta oscillatoria, la metrica poetica può essere interpretata come una forma di **risonanza forzata**: il testo poetico induce nel sistema oscillatorio corticale dell'ascoltatore una sincronizzazione più profonda e più stabile di quella che il parlato naturale produce spontaneamente. Quando il ritmo del verso corrisponde alla frequenza naturale degli oscillatori corticali sillabici (theta), il verso *entra in risonanza* con il sistema: l'accoppiamento tra oscillatore endogeno e stimolo esterno è ottimale, la predizione è massimamente confermata, e — crucialmente — la violazione del metro produce un segnale di errore di previsione particolarmente netto e rilevante.

La sensazione soggettiva di *attesa metricamente guidata* — l'aspettativa che dopo una certa sequenza ritmica segua una certa struttura — è, in questa prospettiva, precisamente la protenzione husserliana realizzata a livello di un sistema oscillatorio entrained al metro del testo. Il piacere estetico della poesia ben costruita potrebbe avere almeno in parte una base oscillatoria: è il piacere della predizione confermata, della risonanza ottimale tra l'oscillazione endogena e quella indotta dal testo. E la violazione deliberata del metro — la *licenza poetica* nei suoi esempi più sofisticati — produce un segnale di errore che non è semplicemente spiacevole ma informativo, portatore di significato pragmatico e stilistico.

Questo argomento ha conseguenze non banali per la teoria della letteratura. Se la risposta estetica al ritmo poetico è parzialmente spiegabile in termini di dinamica oscillatoria, allora la distinzione tradizionale tra un'analisi formale del metro (come struttura astratta di piedi e sillabe) e un'analisi cognitiva della risposta del lettore (come esperienza temporale soggettiva) comincia a dissolversi. La struttura metrica astratta è la struttura di attivazione del sistema oscillatorio corticale; l'esperienza temporale del verso è la sincronizzazione di quel sistema con l'input ritmico. Non vi è un divario tra la struttura e l'esperienza: vi è una singola dinamica che ha una descrizione sia formale che fenomenologica, e la proposta oscillatoria è la cornice che rende visibile l'identità tra le due descrizioni.

XXIV. Il bambino che oscilla: sviluppo preverbale e sincronizzazione primordiale

Se la competenza grammaticale è la struttura degli attrattori del sistema dinamico corticale, e se questa struttura viene plasmata dall'esperienza linguistica nel corso dello sviluppo, allora la domanda sull'acquisizione del linguaggio diventa la domanda su come il cervello neonatale — dotato di un'architettura oscillatoria già presente ma ancora non specializzata per il linguaggio — si trasformi progressivamente nel cervello del parlante adulto, con la sua gerarchia delta-theta-beta-gamma specificamente accoppiata alla struttura della lingua ambientale.

Questa domanda è al tempo stesso empirica e teorica. Empirica, perché richiede dati elettrofisiologici longitudinali su neonati, infanti e bambini nelle varie fasi dell'acquisizione — dati che l'EEG neonatale comincia a fornire, con la qualità e la risoluzione temporale necessarie, solo nei laboratori più avanzati. Teorica, perché richiede un modello del processo di sviluppo che sia compatibile sia con la ricchezza dell'acquisizione naturale che con i vincoli imposti dalla biologia neurale.

Il fatto che i neonati siano già in grado di discriminare la loro lingua materna da lingue straniere — e persino di preferire la voce della madre — sin dalle prime ore di vita è noto dall'epoca storica delle ricerche di DeCasper e collaboratori negli anni Ottanta. Quello che è emerso più di recente, con la disponibilità di paradigmi EEG adatti ai neonati, è che questa competenza percettiva si basa in parte su una sensibilità oscillatoria già presente alla nascita. Il cervello neonatale mostra una sincronia theta con il segnale del parlato — la stessa sincronia sillabica documentata negli adulti — benché meno stabile e meno precisa. La gerarchia oscillatoria è dunque presente, nella sua forma embrionale, sin dai primissimi giorni di vita extrauterina, e probabilmente già durante lo sviluppo fetale, quando il feto è esposto alle strutture prosodiche filtrate della voce materna attraverso il liquido amniotico.

Questo fatto ha implicazioni teoriche considerevoli. Suggerisce che la gerarchia oscillatoria non è interamente il prodotto dell'esperienza linguistica: ha radici biologiche che precedono l'acquisizione. Il cervello arriva all'esperienza linguistica già dotato di una struttura di accoppiamento cross-frequenza che lo predispone a sincronizzarsi con il ritmo del parlato. L'acquisizione non costruisce questa struttura dal nulla: la *sintonizza* progressivamente verso i parametri specifici della lingua ambientale.

XXV. La statistica distributiva come modulazione del paesaggio energetico

La questione centrale dell'acquisizione — come il bambino impari la struttura grammaticale della sua lingua a partire dall'input linguistico senza istruzione esplicita — è stata tradizionalmente divisa tra due posizioni: il **nativismo** (la grammatica è innata, l'acquisizione è semplice messa a punto di parametri preimpostati) e l'**empirismo** (la grammatica si costruisce dall'esperienza, attraverso meccanismi di apprendimento statistico). La proposta oscillatoria offre una terza via che riassume entrambe le tradizioni senza ridursi a nessuna delle due.

Nell'ambito dell'apprendimento statistico, il lavoro seminale di Saffran e collaboratori — pubblicato nel 1996 su *Science* — ha mostrato che infanti di otto mesi sono in grado di segmentare sequenze di sillabe in unità lessicali sulla base delle sole probabilità di transizione tra sillabe: le sillabe che si trovano tipicamente all'interno della stessa parola sono precedute e seguite dalle sillabe vicine con probabilità di transizione elevata, mentre i confini di parola corrispondono a bruschi cali di questa probabilità. Gli infanti imparano questa statistica distribuita senza alcuna istruzione esplicita, dopo soli due minuti di esposizione a sequenze artificiali di sillabe.

In termini oscillatori, l'apprendimento statistico distribuito può essere formalizzato come un processo di aggiornamento del **modello generativo** corticale — precisamente il processo di minimizzazione dell'energia libera discusso nella parte precedente — nel quale le probabilità di transizione vengono codificate come predizioni della gerarchia theta (che anticipa la struttura sillabica) e le loro violazioni vengono segnalate come errori di previsione gamma. L'apprendimento corrisponde alla modifica delle connessioni sinaptiche in modo da ridurre l'energia libera media: i confini di parola, che producono i maggiori errori di previsione (picchi di gamma), vengono progressivamente marcati da transizioni di fase nell'oscillazione theta, realizzando così la segmentazione lessicale come proprietà del sistema oscillatorio.

Formalmente, si può descrivere questo processo come un aggiornamento delle matrici di connessione W tra gli oscillatori del sistema. Se si modella il sistema come una rete di oscillatori Kuramoto con connessioni pesate:

$$d\theta_i/dt = \omega_i + \sum_j W_{ij} \cdot \sin(\theta_j - \theta_i)$$

allora la regola di apprendimento che minimizza l'energia libera predetti corrisponde a un aggiornamento hebbiano generalizzato delle connessioni:

$$\Delta W_{ij} \propto -\partial F / \partial W_{ij} = \langle \epsilon_i \cdot \partial \theta_j / \partial W_{ij} \rangle$$

dove ϵ_i è l'errore di previsione all'oscillatore i e la media è presa su un intervallo di tempo sufficiente. Questa regola di apprendimento ha la struttura di un aggiornamento di gradiente discendente sull'energia libera, e corrisponde, nel dominio neurale, a forme di plasticità sinaptica dipendente dal timing degli spike (*Spike-Timing Dependent Plasticity*, STDP), che sono note per essere un meccanismo fondamentale della plasticità corticale a lungo termine.

L'importanza di questo risultato è che esso collega tre livelli di descrizione che erano stati tradizionalmente mantenuti separati: il livello computazionale dell'apprendimento statistico (le probabilità di transizione), il livello algoritmico della minimizzazione dell'energia libera (l'aggiornamento del modello generativo), e il livello implementazionale della plasticità sinaptica oscillatoriamente guidata (la STDP nelle reti theta-gamma). L'acquisizione del lessico è lo stesso fenomeno descritto a tre livelli di granularità.

XXVI. L'acquisizione della sintassi: oltre la statistica di sequenza

L'apprendimento delle probabilità di transizione tra sillabe spiega la segmentazione lessicale, ma non spiega l'acquisizione della struttura sintattica. Le dipendenze grammaticali non sono statistiche locali di sequenza: sono relazioni strutturali che si estendono su distanze arbitrariamente lunghe nella stringa — e le statistiche locali di sequenza non le catturano. Si può avere la stringa "*L'uomo che tu pensi che sapeva già tutto aveva capito il problema*", in cui la relazione soggetto-verbo principale (*l'uomo ... aveva capito*) attraversa una proposizione relativa incassata di otto parole. Nessuna statistica di adiacenza tra sillabe o tra parole può catturare questa dipendenza.

Questo fatto è stato utilizzato tradizionalmente come argomento in favore del nativismo: se l'apprendimento statistico non è sufficiente per acquisire le dipendenze a lungo raggio della sintassi, allora la struttura che le rende possibili deve essere innata. Tuttavia, il ragionamento trascura una distinzione fondamentale: la distinzione tra statistiche di sequenza *locali* — quelle che Saffran aveva utilizzato nei suoi esperimenti — e statistiche di *struttura*, che catturano la co-occorrenza di elementi non adiacenti mediata dalla loro posizione in una struttura gerarchica.

Nel dominio oscillatorio, la distinzione tra statistiche locali e strutturali corrisponde alla distinzione tra l'accoppiamento a corto raggio (tra oscillatori vicini nella rete corticale) e l'accoppiamento a lungo raggio (tra oscillatori distanti, tipicamente realizzato dalle bande beta). L'acquisizione delle dipendenze a lungo raggio non richiede statistiche di sequenza su finestre ampie: richiede l'apprendimento di quale oscillatore beta deve collegarsi a quale altro oscillatore beta per realizzare la dipendenza strutturale rilevante. Questo apprendimento è mediato da meccanismi di plasticità hebbiana nelle connessioni a lungo raggio — le connessioni del fascicolo arcuato, del fascicolo uncinato, del cingolo — che si consolidano durante l'esposizione a strutture con dipendenze a lungo raggio nella lingua ambientale.

Il bambino che acquisisce l'italiano impara, attraverso la plasticità delle sue connessioni beta, che il soggetto della frase e il verbo principale devono essere collegati oscillatoriamente anche quando sono separati da proposizioni incassate; il bambino che acquisisce il giapponese — una lingua head-final in cui il verbo principale viene sempre alla fine della frase, dopo tutti i suoi argomenti e modificatori — impara a mantenere la connessione beta attiva per un intervallo temporale strutturalmente più lungo prima che il verbo appaia.

Questa differenza nell'organizzazione delle dipendenze beta ha conseguenze misurabili: gli studi di memoria di lavoro sintattica mostrano che i parlanti di lingue a verbo finale hanno una maggiore capacità di mantenimento delle dipendenze strutturali a lungo raggio, compatibile con una connettività beta a lungo raggio più robusta o più efficiente. La fisiologia corticale del parlante di giapponese adulto riflette, nella geometria delle sue connessioni beta, la struttura tipologica della sua lingua — non come una differenza di *hardware* neurologico, ma come una differenza di *settaggio* dell'architettura oscillatoria comune.

XXVII. Il periodo critico come finestra di plasticità oscillatoria

L'esistenza di un **periodo critico** per l'acquisizione del linguaggio — un intervallo temporale dello sviluppo entro il quale l'esposizione alla lingua è necessaria per la completa acquisizione della competenza grammaticale — è uno dei fatti più robustamente documentati della psicologia dello sviluppo, confermato sia da studi su bambini sordi che acquisiscono la lingua dei segni tardivamente, sia da casi tragici di bambini privati di input linguistico nei primi anni di vita, sia da studi sulla difficoltà dell'acquisizione di una seconda lingua nell'età adulta.

Nella prospettiva oscillatoria, il periodo critico può essere reinterpretato come la finestra temporale dello sviluppo in cui la plasticità delle connessioni oscillatorie — e in particolare delle connessioni beta a lungo raggio — è massimamente elevata, permettendo al sistema di essere sintonizzato dall'esperienza linguistica verso i parametri specifici della lingua ambientale. Al di fuori di questa finestra, la

plasticità si riduce — non perché le connessioni diventino permanentemente fisse, ma perché il paesaggio energetico si approfondisce: i bacini di attrazione delle strutture grammaticali acquisite diventano più profondi e le barriere di potenziale tra attrattori diversi diventano più alte, rendendo più difficile per l'esperienza successiva modificare il paesaggio in modo significativo.

Biofisicamente, la chiusura del periodo critico è associata alla maturazione dei sistemi inibitori corticali — in particolare all'incremento dell'espressione della parvalbumina negli interneuroni inibitori a canestro, la cui attività è il principale motore delle oscillazioni gamma locali. Le **reti perineuronali** (*perineuronal nets*), strutture di matrice extracellulare che avvolgono questi interneuroni, si consolidano progressivamente durante lo sviluppo postnatale, riducendo la plasticità sinaptica dei circuiti locali. La maturazione delle reti perineuronali marca la chiusura del periodo critico non soltanto per il linguaggio ma per molte altre funzioni corticali — la visione binoculare, l'elaborazione uditiva — suggerendo che si tratti di un meccanismo plastico generale piuttosto che di un meccanismo specifico per il linguaggio.

In termini formali, la plasticità oscillatoria durante il periodo critico può essere modellata come una variazione temporale della costante di apprendimento nelle regole di aggiornamento hebbiano:

$$\Delta W_{\{ij\}}(t) \propto \lambda(t) \cdot \langle \varepsilon_i \cdot \partial \theta_j / \partial W_{\{ij\}} \rangle$$

dove $\lambda(t)$ è la funzione di plasticità, che ha la forma di una finestra temporale con aumento rapido durante i primi mesi di vita, picco nella fase di massima acquisizione lessicale e grammaticale (2–5 anni), e riduzione progressiva verso la pubertà. La forma precisa di $\lambda(t)$ — la sua larghezza, il suo picco, la sua pendenza di discesa — varia tra le diverse strutture corticali e dipende dal profilo di maturazione dei sistemi inibitori locali.

Questa formalizzazione suggerisce che il periodo critico non sia una soglia discreta — un interruttore che si apre e poi si chiude — ma un processo continuo di graduale riduzione della plasticità. Coerentemente, l'evidenza empirica mostra che l'acquisizione di una seconda lingua è progressivamente più difficile man mano che aumenta l'età di prima esposizione, senza una soglia nettamente discontinua. Il periodo critico è la regione dello spazio (età $\times$ esposizione) in cui $\lambda(t)$ è abbastanza grande da permettere una ristrutturazione significativa del paesaggio energetico grammaticale — e questa regione ha confini sfumati, non netti.

XXVIII. L'innatismo rivisitato: struttura biologica come prior gerarchico

La proposta oscillatoria permette di riformulare la disputa tra nativismo ed empirismo in termini che, finalmente, la rendono empiricamente risolvibile. Aniché chiedersi se la grammatica universale sia innata o appresa — una domanda che, nella sua formulazione tradizionale, oscilla tra alternative incommensurabili — si può chiedere: quale porzione dell'architettura oscillatoria è biologicamente prespecificata e quale è plasmata dall'esperienza?

La risposta che i dati suggeriscono è la seguente. L'architettura gerarchica della gerarchia delta-theta-beta-gamma e la sua struttura di accoppiamento cross-frequenza sono **prior biologici** del sistema: sono prespecificate geneticamente — attraverso i meccanismi di maturazione corticale, il profilo di espressione genica degli interneuroni parvalbuminergici, la geometria anatomica delle connessioni intracorticali a lungo raggio — e si manifestano nelle oscillazioni corticali del neonato, che è già sincronizzato con il ritmo sillabico del parlato. Questo prior biologico corrisponde, nella terminologia chomskyana, alla **Facoltà del Linguaggio** nel suo senso largo (*FLB*, Faculty of Language Broad): la dotazione biologica che permette l'acquisizione di qualunque lingua umana.

Quello che l'esperienza contribuisce è il **settaggio dei parametri** di questa architettura: le frequenze precise di sintonizzazione degli oscillatori per la lingua specifica, la geometria delle connessioni beta per le dipendenze strutturali caratteristiche della lingua ambientale, la presenza o assenza dello strato di accoppiamento tonale per le lingue tonali, la profondità dei bacini di attrazione per le costruzioni sintattiche della lingua acquisita. Questo settaggio corrisponde, nella terminologia chomskyana, alla variazione parametrica che distingue le grammatiche delle singole lingue all'interno dell'insieme universalmente possibile.

L'innatismo e l'empirismo non sono dunque in conflitto: descrivono aspetti complementari dello stesso sistema. L'innato è il prior oscillatorio; l'acquisito è la sua modulazione attraverso l'esperienza. La grammatica universale è la struttura biologicamente prespecificata della gerarchia oscillatoria corticale; la grammatica della singola lingua è il settaggio specifico di quella struttura che l'esperienza determina entro la finestra del periodo critico. E il problema dell'acquisizione del linguaggio — il tradizionale *poverty of the stimulus argument* — si dissolve non perché l'input sia più ricco di quanto Chomsky pensasse, ma perché il sistema biologico che lo elabora è già strutturato in modo da renderlo sufficiente: un sistema con i giusti prior biologici non ha bisogno di un input illimitatamente ricco per convergere verso la grammatica corretta.

XXIX. L'afasia come geometria distorta

Se la competenza grammaticale è la struttura degli attrattori del sistema dinamico corticale, allora le patologie del linguaggio — le afasie, le disfasie, le anomalie dell'elaborazione sintattica nelle psicopatologie — non sono semplici perdite di funzioni localizzate: sono perturbazioni della geometria del paesaggio energetico

grammaticale. Questa formulazione non è soltanto metaforica: ha implicazioni diagnostiche e terapeutiche concrete, e permette di leggere la fenomenologia clinica delle afasie in una luce nuova.

L'afasia di Broca — la forma più studiata, prodotta da lesioni del giro frontale inferiore sinistro e delle regioni corticali e sottocorticali circostanti — è tradizionalmente descritta come caratterizzata da produzione ridotta e sforziosa, prosodia appiattita, prevalenza di parole di contenuto rispetto alle parole grammaticali (preposizioni, congiunzioni, morfemi flessivi), e comprensione relativamente preservata per le frasi semplici ma compromessa per le frasi con struttura sintattica complessa — in particolare quelle con dipendenze a lungo raggio, come le frasi relative oggetto e le costruzioni passive non canoniche. La difficoltà con le relative oggetto — *"Il bambino che la maestra ha lodato ha vinto il premio"*, in cui il referente del pronome relativo è l'oggetto del verbo incassato — è una delle firme più caratteristiche dell'afasia di Broca, ed è rimasta storicamente difficile da spiegare in termini puramente lesionali.

Nella prospettiva oscillatoria, la difficoltà con le dipendenze a lungo raggio in Broca riflette una perturbazione specifica delle connessioni beta a lungo raggio che realizzano il mantenimento della traccia dell'elemento estratto durante l'elaborazione del materiale intermedio. L'area di Broca — il giro frontale inferiore sinistro — è un nodo centrale nelle reti frontale-temporali che mediano l'oscillazione beta sintattica: la sua lesione non abolisce il sistema oscillatorio ma ne altera la geometria delle connessioni, riducendo la profondità del bacino di attrazione per le strutture con dipendenze a lungo raggio. Queste strutture non diventano completamente inaccessibili, ma richiedono un investimento computazionale molto maggiore — un segnale di errore di previsione più elevato, una traiettoria più lunga nello spazio degli stati — per convergere verso l'attrattore corretto.

Questa previsione è testabile: ci si aspetta che i pazienti con afasia di Broca mostrino una riduzione del PAC theta-beta e del PLV beta durante l'elaborazione di frasi con dipendenze a lungo raggio, con una correlazione negativa tra la gravità del deficit sintattico e il grado di accoppiamento beta nelle regioni frontali-temporali sinistre. Le evidenze disponibili sono compatibili con questa previsione, benché studi EEG di alta qualità su popolazioni afasiche sufficientemente ampie siano ancora scarsi.

XXX. La dislessia come dissincronia sillabica

La **dislessia evolutiva** — la difficoltà specifica nell'apprendimento della lettura che interessa circa il 5–10% della popolazione in età scolare — è stata oggetto di decenni di ricerca neuropsicologica che ha prodotto numerose teorie parziali (deficit fonologico, deficit magnocellulare, deficit cerebellare) senza convergere su un quadro esplicativo unitario. La prospettiva oscillatoria offre una proposta

unificante che, pur non risolvendo tutti i problemi aperti, fornisce un framework coerente per integrare le evidenze disponibili.

La tesi centrale, elaborata in modo indipendente da Goswami e collaboratori e da Giraud e Ramus, è che la dislessia sia fondamentalmente una perturbazione della **sincronia sillabica theta**: il sistema oscillatorio dei dislessici mostra una ridotta capacità di sincronizzarsi con le strutture temporali del parlato alla frequenza theta, compromettendo la segmentazione sillabica che è il prerequisito per la codifica fonologica delle parole scritte.

La misurazione di questa dissincronia si basa sulla quantificazione della **coerenza** tra il segnale acustico del parlato (o della musica ritmica) e le oscillazioni corticali theta dell'ascoltatore, attraverso l'indice di coerenza cortico-acustica:

$$\text{Coh}(f) = |\langle X^*(f) \cdot Y(f) \rangle|^2 / (\langle |X(f)|^2 \rangle \cdot \langle |Y(f)|^2 \rangle)$$

dove $X(f)$ è lo spettro di potenza del segnale acustico, $Y(f)$ è lo spettro di potenza del segnale EEG, e la media è presa su trial multipli. La coerenza $\text{Coh}(f)$ è un numero reale nell'intervallo $[0, 1]$: vale 0 in assenza di relazione tra i due segnali e 1 in caso di relazione lineare perfetta. Nei soggetti tipici, la coerenza cortico-acustica mostra un picco pronunciato nella banda theta (4–8 Hz) durante l'ascolto di parlato naturale; nei soggetti dislessici, questo picco è sistematicamente ridotto — a volte assente — nella banda theta, mentre non sono presenti differenze significative nelle bande più alte (beta, gamma).

Questo deficit di sincronia theta ha una conseguenza diretta sull'acquisizione della lettura. Imparare a leggere in un sistema alfabetico richiede di stabilire corrispondenze sistematiche tra i grafemi (le lettere) e i fonemi (i suoni) della lingua. Per fare questo, il bambino deve essere in grado di *segmentare* le parole orali in unità fonologiche discrete — un'operazione che richiede la capacità di marcare i confini sillabici e, all'interno delle sillabe, i confini tra fonemi. Nella prospettiva oscillatoria, questa segmentazione è realizzata dalle transizioni di fase theta (per i confini sillabici) e dalle oscillazioni gamma annidate entro ciascun ciclo theta (per i confini fonemici). Se la sincronia theta è distorta, la segmentazione sillabica è incerta, e le corrispondenze grafema-fonema non possono essere acquisite con l'automatismo e la precisione necessari.

Crucialmente, questa teoria predice che il deficit dei dislessici sia nella dominio temporale del processamento fonologico, non nella struttura astratta delle rappresentazioni fonologiche in sé. I dislessici non hanno rappresentazioni fonologiche errate o impoverite in senso categoriale: hanno una difficoltà nel sincronizzare le proprie oscillazioni theta con la struttura temporale del parlato, il che rende meno nitida e meno stabile la segmentazione sillabica. La previsione più specifica — che interventi ritmici e musicali, che allenano la sincronizzazione theta con stimoli ritmici regolari, migliorino le competenze di decodifica

fonologica nei dislessici — ha ricevuto alcune conferme sperimentali, ancora preliminari ma sufficientemente robuste da giustificare ulteriori indagini.

XXXI. La schizofrenia e il disaccoppiamento gamma-theta

Le **psicosi schizofreniche** sono caratterizzate, sul piano cognitivo, da un insieme di deficit che includono alterazioni del pensiero formale (illogicità, derailment, tangenzialità), difficoltà nell'elaborazione del linguaggio (comprensione ridotta di metafore, ironia, discorso indiretto), e perturbazioni della memoria di lavoro e dell'attenzione. L'analisi neurofisiologica della schizofrenia ha identificato una firma oscillatoria caratteristica: una **riduzione delle oscillazioni gamma** nelle aree prefrontali e temporali, associata a una riduzione del PAC theta-gamma.

La base biofisiologica di questo deficit è relativamente ben compresa: la schizofrenia è associata a una riduzione dell'espressione della parvalbumina negli interneuroni inibitori a canestro delle aree prefrontali e temporali — precisamente le cellule che generano le oscillazioni gamma locali. La riduzione di questi interneuroni compromette la generazione locale di gamma, e — per il meccanismo di accoppiamento bottom-up di cui si è discusso — riduce anche la precisione con cui i segnali di errore di previsione (trasportati dal gamma) vengono trasmessi verso i livelli superiori della gerarchia predittiva.

In termini del framework della codifica predittiva, questo si traduce in una riduzione della **precisione degli errori di previsione**: i segnali di errore ascendenti sono più deboli e più rumorosi, il che significa che il sistema aggiorna le proprie predizioni in modo meno efficiente. Le predizioni dei livelli superiori — i prior semantici e sintattici — hanno dunque un peso relativo maggiore rispetto ai segnali di errore, producendo un sistema che è dominato dai propri prior anziché essere corretto dall'input. Questa dinamica è precisamente quella che Friston ha proposto come modello computazionale delle psicosi: un sistema in cui la precisione relativa di prior ed errori di previsione è alterata, con un eccesso di peso dei prior che porta a percepire pattern anche nell'assenza (allucinazioni) e a mantenere credenze rigide contro l'evidenza (deliri).

La conseguenza linguistica di questo meccanismo è una compromissione selettiva dell'elaborazione degli aspetti del linguaggio che richiedono un aggiornamento frequente e preciso delle predizioni in funzione dell'input: l'ironia (che richiede di aggiornare la rappresentazione pragmatica alla luce di segnali contestuali sottili), le metafore nuove (che richiedono di integrare significati letterali con proiezioni concettuali inattese), il discorso a lungo raggio (che richiede di mantenere aggiornato il modello del discorso attraverso molti cicli di aggiornamento). La schizofasia — il pensiero formalmente disorganizzato del paziente schizofrenico — può essere letta come il linguaggio di un sistema la cui gerarchia predittiva è sistematicamente distorta dalla riduzione della precisione

gamma: un sistema che predilige le proprie associazioni interne rispetto alle aspettative strutturali dell'interlocutore, perché i segnali di errore che dovrebbero correggere queste associazioni sono attenuati dal deficit dell'interneurone parvalbuminergico.

XXXII. L'autismo e la iper-precisione delle previsioni locali

Un contrasto illuminante con la schizofrenia è offerto dal **disturbo dello spettro autistico** (DSA), che presenta un profilo cognitivo e neurofisiologico in molti rispetti speculare. Se la schizofrenia si caratterizza per una riduzione della precisione degli errori di previsione gamma (con eccesso di peso dei prior), il DSA si caratterizza — secondo la teoria di Friston e Lawson, e di Pellicano e Burr — per una condizione di iper-precisione dei segnali sensoriali, con una difficoltà a soppesare opportunamente i prior contestuali rispetto all'input immediato.

Neurofisiologicamente, il DSA è associato non a una riduzione delle oscillazioni gamma ma a un **eccesso di oscillazioni gamma locali** nelle aree sensoriali e associative, con una ridotta capacità di integrare questi segnali locali nella gerarchia delta-theta delle scale temporali più lunghe. Il rapporto tra l'elaborazione locale ad alta risoluzione (gamma) e l'organizzazione globale a bassa risoluzione (delta-theta) è alterato in senso opposto rispetto alla schizofrenia: il sistema è iper-sensibile ai dettagli locali ma meno capace di integrarli in strutture sovraordinate.

Nel dominio linguistico, questo si manifesta come una dissociazione tra competenza grammaticale formale — spesso preservata, e talvolta superiore alla media nelle aree relative alla grammatica regolare e alla morfologia paradigmatica — e comprensione pragmatica e discorsiva, sistematicamente compromessa. I bambini con DSA possono acquisire una grammatica formalmente accurata ma faticano con l'ironia, le implicature conversazionali griciane, la gestione del referente attraverso il discorso — tutti aspetti del linguaggio che richiedono l'integrazione di segnali locali con prior globali di natura pragmatica e discorsiva.

La proposta oscillatoria traduce questa dissociazione clinica in termini di architettura della gerarchia di accoppiamento: nel DSA, l'accoppiamento gamma-locale è eccessivamente forte e relativamente autonomo dalla modulazione delta-theta — la gerarchia non è distrutta ma è squilibrata verso il basso, con l'elaborazione locale che non viene sufficientemente vincolata dall'impalcatura temporale delle scale più lente. Il termine clinico di iper-sensorialità — la percezione sopraffacente di stimoli sensoriali che molti soggetti con DSA riferiscono — ha il suo correlato neurofisiologico precisamente in questo eccesso di gamma locale non adeguatamente modulato dalla gerarchia superiore.

Considerati insieme, il profilo della schizofrenia (riduzione della precisione gamma, eccesso di peso dei prior) e quello del DSA (eccesso di precisione gamma, riduzione del peso dei prior) suggeriscono che la gerarchia oscillatoria corticale sia un sistema in equilibrio dinamico tra due tendenze opposte: la predizione discendente e la correzione ascendente. La cognizione normale — e il linguaggio normale come sua componente privilegiata — si realizza nell'equilibrio tra queste due tendenze. La patologia si manifesta quando l'equilibrio è perturbato in una direzione o nell'altra — un'immagine che ha la chiarezza e la falsificabilità proprie delle buone teorie scientifiche.

XXXIII. Il problema della discontinuità filogenetica

Tra tutti i problemi che circondano il linguaggio umano, quello filogenetico è forse il più refrattario e al tempo stesso il più vertiginoso: perché soltanto gli esseri umani posseggono una sintassi ricorsiva, mentre i loro parenti filogeneticamente più prossimi — gli scimpanzé, i bonobo, i gorilla — non ne mostrano alcun segno nonostante decenni di tentativi di insegnamento? Il distacco genetico tra *Homo sapiens* e *Pan troglodytes* è di appena il 1.2–1.6% del genoma; eppure questo minuscolo intervallo separa un sistema cognitivo che non produce spontaneamente struttura sintattica da uno che la produce obbligatoriamente, automaticamente, e in modo illimitatamente generativo.

Chomsky, Fitch e Hauser, nel celebre articolo del 2002 su *Science* — "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" — avanzano la proposta che la peculiarità del linguaggio umano risieda nella ricorsività, e che la ricorsività sia una novità evolutiva relativamente recente e specifica della nostra specie. Nella versione più radicale di questa posizione, Merge — l'operazione che genera struttura gerarchica — è l'unica componente della Facoltà del Linguaggio nel suo senso stretto (*FLN*, Faculty of Language Narrow) che è esclusivamente umana: tutto il resto — le capacità vocali, la memoria associativa, la teoria della mente — è condiviso, in varia misura, con altri animali.

La proposta oscillatoria aggiunge una dimensione nuova a questo dibattito. Se Merge è implementato dall'accoppiamento cross-frequenza tra oscillazioni neurali, allora la domanda evoluzionistica diventa: quale mutazione o graduale riarrangiamento della circuiteria corticale ha reso possibile il tipo di accoppiamento gerarchico che realizza la ricorsività? E questa domanda è in linea di principio empiricamente accessibile, perché l'architettura oscillatoria del cervello dei primati non umani è studiabile con le stesse tecniche — elettrofisiologia, EEG, MEG — che si usano per i parlanti umani.

XXXIV. Il coupling cross-frequenza nei primati non umani

Le oscillazioni neurali sono presenti e studiate nei primati non umani — in particolare nei macachi — da almeno quattro decenni. Gli studi di elettrofisiologia invasiva su macachi hanno documentato oscillazioni nella corteccia prefrontale, nel giro temporale superiore e nelle aree motorie con caratteristiche simili a quelle umane: gamma locali generati dai circuiti di interneuroni parvalbuminergici, beta di mantenimento nelle aree prefrontali, delta-theta nelle corteccie temporali. L'accoppiamento cross-frequenza — e in particolare il PAC theta-gamma nelle aree prefrontali — è stato documentato anche nei macachi, soprattutto in relazione a compiti di memoria di lavoro.

Ciò che sembra assente, o almeno drasticamente ridotto, nei primati non umani è la struttura *gerarchica* di questo accoppiamento: la modulazione del theta da parte del delta, e la conseguente organizzazione su tre o più livelli temporali. Nei macachi, l'accoppiamento PAC è documentato principalmente per la coppia theta-gamma, ma non vi è evidenza robusta di un accoppiamento delta-theta-gamma a tre livelli che organizzi l'elaborazione su scale temporali corrispondenti a quelle sintattiche. In altri termini, la gerarchia oscillatoria dei macachi sembra arrestarsi a due livelli (theta-gamma), mentre quella umana si estende almeno a tre (delta-theta-gamma) e probabilmente a quattro (infra-delta-delta-theta-gamma) nel dominio discorsivo.

Questa differenza — se confermata da studi sistematici con protocolli comparabili — avrebbe un'implicazione teorica di grande portata: la ricorsività sintattica, implementata da Merge come operazione che si applica ai propri output, richiederebbe un accoppiamento oscillatorio gerarchico a più livelli che i primati non umani non possiedono. La comparsa della ricorsività nella linea evolutiva umana coinciderebbe con la comparsa di un terzo livello di accoppiamento — la modulazione del ritmo theta da parte del ritmo delta — che crea la finestra temporale necessaria per elaborare strutture sintattiche di frase intera come unità, annidate entro strutture discorsive più ampie.

XXXV. La genetica delle oscillazioni: FOXP2 e le reti per il ritmo

Il gene **FOXP2** — il primo gene associato a un deficit specifico del linguaggio, identificato nei primi anni Duemila nello studio della famiglia KE da Fisher, Vargha-Khadem e collaboratori — è diventato rapidamente il simbolo della base genetica del linguaggio. I membri della famiglia KE portatori di una mutazione nel gene FOXP2 mostrano un disturbo severo della coordinazione orofacciale e della grammatica, con difficoltà sia nella produzione che nella comprensione di frasi sintatticamente complesse.

Ma FOXP2 non è, come talvolta viene presentato nei resoconti divulgativi, un "gene del linguaggio". È un fattore di trascrizione — una proteina che regola l'espressione di centinaia di altri geni — con un profilo di espressione molto ampio, che include non soltanto le aree corticali del linguaggio ma anche i gangli

della base, il cervelletto, il talamo e l'ippocampo. La sua versione umana differisce da quella dello scimpanzé per due sostituzioni aminoacidiche in posizioni critiche del dominio di legame al DNA — mutazioni che alterano il profilo di geni target e che si sono fissate nella linea evolutiva umana in un momento relativamente recente, probabilmente negli ultimi 200.000–400.000 anni, in coincidenza con la comparsa del linguaggio moderno.

Nei topi e nei cantori — gli uccelli canori, che hanno sviluppato in modo indipendente la capacità di imparare vocalmente — FOXP2 regola la plasticità dei circuiti striatali e cerebellari implicati nell'apprendimento motore delle sequenze vocali. Nei topi con la versione umana di FOXP2 (knock-in), si osservano alterazioni nell'apprendimento di sequenze motorie e, crucialmente, modificazioni nelle proprietà elettrofisiologiche dei neuroni striatali — inclusa una modifica delle loro correnti ioniche che altera la finestra temporale di integrazione.

Nella prospettiva oscillatoria, queste osservazioni suggeriscono che FOXP2 sia coinvolto, tra l'altro, nella determinazione delle proprietà temporali delle reti corticali e sottocorticali che realizzano la gerarchia oscillatoria. La versione umana di FOXP2 potrebbe aver contribuito all'emergenza del terzo livello di accoppiamento — la modulazione delta del ritmo theta — allungando le finestre di integrazione temporale delle reti striatali e talamostriatali che regolano l'oscillazione delta corticale.

Questa proposta è ancora altamente speculativa. La distanza tra la biologia molecolare di un fattore di trascrizione e la neurofisiologia dell'accoppiamento delta-theta è immensa, e i passaggi causali che le collegano sono ancora largamente oscuri. Ma la direzione del ragionamento è fondata: la variazione genetica che distingue *Homo sapiens* da *Pan troglodytes* — incluse le varianti di FOXP2 e delle centinaia di geni che esso regola — ha prodotto differenze nelle proprietà biofisiche dei circuiti corticali che si manifestano, al livello della popolazione, come differenze nell'architettura oscillatoria. E l'architettura oscillatoria, nella proposta qui sostenuta, è il substrato computazionale della ricorsività sintattica.

XXXVI. Il canto degli uccelli e la convergenza evolutiva

Un argomento indipendente e particolarmente suggestivo a favore della proposta oscillatoria viene dalla biologia comparata del **canto degli uccelli**. Gli uccelli canori — i Passeri dell'ordine Oscines — hanno sviluppato in modo evolutivamente indipendente rispetto ai primati la capacità di imparare vocalmente nuovi suoni attraverso l'imitazione e la pratica, una capacità che condividono con i cetacei, i pipistrelli e alcuni pinnipedi, ma che è assente nella maggior parte dei mammiferi non umani inclusi gli altri primati.

Il canto degli uccelli ha struttura gerarchica: le frasi (sequenze di note prodotte in un'unica espirazione) sono composte da sillabe, le sillabe da note, le note da elementi sonori elementari. Questa gerarchia è imparata durante il **periodo critico** del canto — un intervallo dello sviluppo giovanile in cui l'uccello ascolta e memorizza il canto del suo tutore, per poi riprodurlo attraverso un processo di pratica guidata dal feedback. Le aree cerebrali che controllano il canto negli uccelli — l'Area X, l'HVC, il RA — formano circuiti la cui architettura ricorda sorprendentemente i circuiti corticali-striatali umani per il linguaggio, benché gli uccelli siano filogeneticamente lontanissimi dai primati.

La domanda oscillatoria è: l'HVC e le aree connesse mostrano accoppiamento cross-frequenza gerarchico durante la produzione e l'apprendimento del canto? La risposta disponibile è affermativa: nei zebrafanchi (*Taeniopygia guttata*), l'area HVC mostra oscillazioni theta durante la produzione del canto, e queste oscillazioni mostrano accoppiamento con oscillazioni gamma locali in modo compatibile con una struttura di elaborazione gerarchica. L'HVC del zebrafanco è, nei fatti, un orologio theta che temporizza la produzione sequenziale delle sillabe del canto: ciascun ciclo theta corrisponde approssimativamente alla durata di una sillaba, e le popolazioni di neuroni HVC scaricano in sequenza, ciascuna durante una fase specifica del ciclo theta, producendo la sequenza ordinata delle sillabe.

Questo risultato è di importanza teorica enorme, perché suggerisce che la sincronia theta come meccanismo di temporizzazione delle sequenze vocali sia una soluzione evolutiva convergente — scoperta indipendentemente negli uccelli e nei primati — al problema computazionale di produrre e apprendere sequenze di unità discrete con struttura gerarchica. La convergenza evolutiva è uno degli argomenti più forti per l'esistenza di un meccanismo: se sistemi biologici indipendenti convergono sulla stessa soluzione, è perché quella soluzione è particolarmente efficace rispetto all'alternativa che il panorama delle possibilità biologiche offre.

La differenza cruciale tra il canto degli uccelli e il linguaggio umano rimane comunque la ricorsività: il canto degli uccelli ha struttura gerarchica, ma la gerarchia è fissa e finita — non vi è nel canto degli uccelli l'equivalente di Merge, che genera strutture illimitate a partire da strutture finite. Il canto ha la struttura di una grammatica finita a stati — un automa finito, nella terminologia di Chomsky — non di una grammatica libera da contesto. Questo è esattamente il limite che l'accoppiamento a due livelli (theta-gamma nell'HVC) riflette: una gerarchia con profondità limitata, non illimitatamente ricorsiva. La ricorsività illimitata richiederebbe l'estensione della gerarchia oscillatoria a un terzo livello (delta-theta-gamma), che gli uccelli non sembrano possedere — un'assenza che corrisponde precisamente all'assenza della ricorsività nel loro sistema di comunicazione.

L'evoluzione della mente linguistica può dunque essere raccontata, in questa prospettiva, come l'evoluzione da una gerarchia oscillatoria a due livelli —

presente, in varie forme, in molti vertebrati con capacità di apprendimento vocale — a una gerarchia a tre o più livelli, esclusiva della linea evolutiva umana, che abilita la ricorsività e con essa la capacità di generare strutture sintattiche di complessità illimitata. La storia dell'origine del linguaggio non è la storia dell'emergenza improvvisa di un organo radicalmente nuovo: è la storia di un'amplificazione gerarchica di un sistema oscillatorio che aveva già i suoi rudimenti in antenati filogeneticamente distanti.

XXXVII. Il grande ritorno del connessionismo e il trionfo dell'attenzione

Nei decenni compresi tra gli anni Ottanta e Novanta del secolo scorso, il connessionismo — la proposta che i processi cognitivi possano essere modellati come l'attività di reti di unità semplici connesse da pesi sinaptici modificabili — aveva sfidato il paradigma della cognizione come manipolazione di simboli discreti, senza tuttavia riuscire a produrre sistemi che elaborassero struttura linguistica in modo convincente. Le reti ricorrenti di quell'epoca faticavano con le dipendenze a lungo raggio, non generalizzavano adeguatamente alle strutture non viste in addestramento, e non mostravano la sensibilità alle violazioni strutturali che caratterizza i parlanti umani.

Tutto questo è cambiato radicalmente con l'introduzione dei **Transformer** — l'architettura introdotta da Vaswani e collaboratori nel 2017 con il celebre articolo *"Attention Is All You Need"* — e con il successivo sviluppo dei grandi modelli linguistici (GPT, BERT, LLaMA e i loro successori) che su questa architettura si basano. Questi sistemi mostrano capacità linguistiche che, superficialmente, sembrano catturare aspetti significativi della competenza grammaticale umana: rispettano le dipendenze a lungo raggio, producono frasi grammaticalmente corrette, risolvono ambiguità pragmatiche, parafrasano, traducono, e persino argomentano con una coerenza che ha sorpreso molti osservatori.

La domanda che la proposta oscillatoria pone a questi sistemi è radicale: in che senso elaborano struttura, e in che senso il loro modo di elaborarla è rilevante per la comprensione della cognizione linguistica umana? La risposta che si argomenterà è la seguente: i Transformer catturano alcune proprietà statistiche della struttura sintattica in modo sorprendentemente efficace, ma il meccanismo mediante il quale lo fanno è fondamentalmente diverso — per natura e non soltanto per implementazione — dal meccanismo oscillatorio che il cervello umano utilizza. Questa differenza non è un dettaglio di ingegneria: è una differenza teoricamente significativa che illumina per contrasto la specificità della soluzione biologica.

XXXVIII. Il meccanismo dell'attenzione e la sua struttura

Il cuore computazionale dei Transformer è il meccanismo di **self-attention** (auto-attenzione), che permette a ciascuna posizione nella sequenza di input di integrare informazione da tutte le altre posizioni, con pesi che dipendono dalla rilevanza reciproca tra le posizioni. Formalmente, dato un insieme di vettori di input $X = [x_1, \dots, x_n] \in \mathbb{R}^{n \times d}$, il meccanismo di attenzione proietta ciascun vettore in tre spazi distinti attraverso matrici di proiezione apprese:

$$Q = X \cdot W_Q, \quad K = X \cdot W_K, \quad V = X \cdot W_V$$

dove $Q \in \mathbb{R}^{n \times d_k}$ sono le query, $K \in \mathbb{R}^{n \times d_k}$ le chiavi, e $V \in \mathbb{R}^{n \times d_v}$ i valori. L'output dell'attenzione è:

$$\text{Attention}(Q, K, V) = \text{softmax}(Q \cdot K^T / \sqrt{d_k}) \cdot V$$

Il termine $Q \cdot K^T$ è una matrice di similarità $n \times n$ tra tutte le coppie di posizioni: la sua entrata (i, j) misura la rilevanza della posizione j per l'elaborazione della posizione i . La divisione per $\sqrt{d_k}$ serve a stabilizzare i gradienti durante l'addestramento. La funzione softmax normalizza i pesi a somma 1 lungo la dimensione j , producendo una distribuzione di probabilità sull'intera sequenza che determina come le informazioni vengono combinate.

Il meccanismo di **multi-head attention** estende questo schema eseguendo h computazioni di attenzione in parallelo su sottospazi diversi:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \cdot W_O$$

dove $\text{head}_k = \text{Attention}(Q \cdot W_Q^k, K \cdot W_K^k, V \cdot W_V^k)$. L'intuizione è che teste diverse apprendono relazioni diverse tra le posizioni: alcune catturano relazioni sintagmatiche locali (soggetto-verbo adiacente), altre relazioni a lungo raggio (soggetto-verbo separati da una relativa), altre ancora relazioni semantiche (sinonimi, iponimi).

Ciò che questo meccanismo realizza è una forma di **attenzione globale non locale**: ogni posizione può, in un singolo passaggio, integrare informazione da qualunque altra posizione nella sequenza, con pesi appresi dall'addestramento. La dipendenza a lungo raggio tra soggetto e verbo in una frase con relativa incassata non pone nessuna difficoltà particolare al Transformer: le due posizioni si "attraggono" attraverso il meccanismo di attenzione indipendentemente dalla distanza che le separa.

XXXIX. Ciò che i Transformer non hanno: tempo, corpo, mondo

L'efficacia computazionale dei Transformer è incontestabile — i loro risultati su benchmark linguistici sono straordinari — ma nasconde una differenza strutturale fondamentale rispetto al sistema oscillatorio biologico. Questa differenza ha diversi aspetti complementari che conviene distinguere con cura.

Il primo e più ovvio è l'**assenza di temporalità intrinseca**. Il Transformer elabora la sequenza di input come un oggetto statico nello spazio: tutte le posizioni sono disponibili simultaneamente, e l'attenzione integra informazione da tutte in un singolo passaggio (o in un numero fisso di passaggi, pari alla profondità della rete). Non vi è un tempo di elaborazione che si dispiega progressivamente, non vi è una finestra di integrazione che si apre e si chiude, non vi è una struttura di ritenzione-protenzione: c'è soltanto una matrice $n \times n$ di pesi di attenzione che cattura, staticamente, le relazioni tra tutte le coppie di posizioni.

Questa atemporalità ha conseguenze precise. Il Transformer non può, per costruzione, modellare la sensibilità *online* alla struttura della frase — la progressiva costruzione della rappresentazione sintattica mentre la frase si dispiega nell'input. Eppure è questa sensibilità online che gli esperimenti di tempi di reazione e di registrazione EEG misurano nel parlante umano: la P600 emerge circa 600 millisecondi dopo la violazione sintattica, non dopo che tutta la frase è stata letta; la N400 emerge circa 400 millisecondi dopo la parola semanticamente anomala, non dopo la conclusione della frase. Il cervello umano elabora il linguaggio in tempo reale, integrando progressivamente l'input e generando previsioni continue sul materiale futuro. Il Transformer non lo fa.

Il secondo aspetto è l'**assenza di corpo e situazione**. Il modello linguistico è addestrato su testo scritto — sequenze di token — senza alcun accesso al mondo fisico, ai contesti situazionali, agli interlocutori incarnati. Il significato, in questi sistemi, è puramente distribuzionale: emerge dalle co-occorrenze statistiche tra token nel corpus di addestramento. Ma il significato nel senso rilevante per la teoria linguistica — il riferimento, la denotazione, le condizioni di verità — è una relazione tra espressioni linguistiche e stati del mondo, non tra espressioni linguistiche tra di loro.

Nella prospettiva della codifica predittiva incarnata, il modello generativo che il cervello elabora non è soltanto un modello del testo linguistico: è un modello del mondo — degli oggetti, degli eventi, delle relazioni causali, degli agenti e delle loro intenzioni — che viene aggiornato dall'input linguistico ma non si riduce ad esso. Le oscillazioni corticali che realizzano la gerarchia sintattica sono parte di un sistema più ampio che integra input visivi, propriocettivi, emozionali e mnemonici in un modello unificato dell'ambiente. Il linguaggio, in questo sistema, non è elaborazione di token ma *azione sul mondo* — nel senso preciso di Wittgenstein: uno strumento per fare cose, non soltanto per descriverle.

Il terzo aspetto, forse il più sottile, è l'**assenza di apprendimento causale**. I Transformer apprendono correlazioni statistiche tra token, ma non strutture

causali — non il fatto che certe strutture linguistiche siano grammaticali perché generate da un sistema computazionale biologico con certe proprietà, ma il fatto che certi pattern di co-occorrenza siano più frequenti di altri nel corpus. Questa distinzione — tra correlazione statistica e struttura causale generativa — è precisamente ciò che separa l'apprendimento statistico distribuito dall'acquisizione della grammatica nel senso minimalista.

XL. La critica della lunghezza del contesto e la memoria oscillatoria

Un confronto più preciso tra il meccanismo dell'attenzione e il meccanismo oscillatorio rivela una differenza importante nella modalità di gestione della **dipendenza a lungo raggio**. Nel Transformer, la dipendenza tra posizioni lontane è gestita attraverso l'attenzione globale: qualunque due posizioni possono interagire direttamente, con pesi che dipendono dalla loro similarità nel spazio delle query-chiavi. Il costo computazionale di questa operazione è $O(n^2)$ in tempo e spazio, dove n è la lunghezza della sequenza — il che rende i Transformer standard computazionalmente proibitivi per sequenze molto lunghe.

Nel sistema oscillatorio biologico, le dipendenze a lungo raggio sono gestite attraverso un meccanismo radicalmente diverso: le oscillazioni beta mantengono attive le rappresentazioni delle posizioni precedenti rilevanti mentre nuove posizioni vengono elaborate, senza che sia necessario accedere simultaneamente a tutte le posizioni. Il costo computazionale di questo meccanismo scala in modo molto più favorevole: la quantità di informazione mantenuta attiva in beta è limitata dalla capacità della memoria di lavoro — il famoso vincolo delle 7 ± 2 unità di Miller — ma questa limitazione è compensata dall'architettura gerarchica del sistema, che aggrega le posizioni in unità sempre più grandi (sillabe → parole → sintagmi → frasi) prima di mantenerle in beta.

In altri termini, il Transformer mantiene tutto in memoria simultaneamente (a costo $O(n^2)$) mentre il sistema biologico mantiene una rappresentazione gerarchicamente compressa delle posizioni precedenti (a costo lineare in n , ma con perdita di dettaglio alle scale più fini). Questa differenza ha conseguenze sulle proprietà di generalizzazione dei due sistemi: il Transformer può, in linea di principio, catturare qualunque dipendenza all'interno della sua finestra di contesto, ma decade rapidamente al di fuori di essa; il sistema biologico gestisce le dipendenze attraverso la struttura della gerarchia oscillatoria, che è in linea di principio illimitata nella sua profondità ricorsiva ma limitata nella quantità di dettaglio fine che conserva alle scale temporali più brevi.

Questa analisi suggerisce che le due architetture — oscillatoria biologica e attention-based artificiale — siano soluzioni diverse allo stesso problema computazionale, con profili di forze e debolezze complementari. Il Transformer è superiore nella cattura di correlazioni statistiche a lungo raggio su dataset

enormi; il sistema oscillatorio è superiore nell'elaborazione online in tempo reale, nella generalizzazione strutturale a nuove sequenze, e nella capacità di integrare il segnale linguistico con informazioni multimodali e situazionali. Il fatto che i grandi modelli linguistici producano output linguisticamente plausibili non implica che elaborino il linguaggio nello stesso modo in cui lo elabora il cervello umano: implica soltanto che la plausibilità linguistica è una proprietà statisticamente catturabile da correlazioni distribuzionali, il che è un risultato interessante ma non equivalente alla comprensione.

XLI. L'illusione della comprensione e il test della struttura

Vi è un test cruciale che permette di distinguere, almeno in linea di principio, tra sistemi che elaborano struttura genuinamente e sistemi che ne simulano gli effetti attraverso correlazioni statistiche. Il test consiste nel presentare ai sistemi delle strutture **strutturalmente anomale ma statisticamente plausibili** — frasi che non violano le distribuzioni di co-occorrenza statistiche del corpus di addestramento ma violano i vincoli della grammatica formale.

Un esempio classico è la distinzione tra dipendenze **strutturali** e **lineari** nell'accordo soggetto-verbo in presenza di un nome attrattore. Nella frase *"Il libro che gli studenti hanno comprato è esaurito"*, il verbo principale *è* deve concordare con il soggetto *libro* (singolare) e non con il nome più vicino *studenti* (plurale). Un sistema che rispetta il vincolo strutturale produce *è*; un sistema che segue la statistica lineare di prossimità è tentato verso *sono*. I parlanti umani rispettano sistematicamente il vincolo strutturale; i modelli linguistici, testati su questo tipo di materiale, mostrano una tendenza — variabile tra modelli diversi e tra strutture diverse — verso l'errore di attrazione lineare.

Nella prospettiva oscillatoria, la ragione per cui i parlanti umani non commettono questo errore è strutturale: la connessione beta tra il soggetto e il verbo principale è stabilita al momento in cui il sintagma soggetto viene elaborato, e questa connessione viene mantenuta attiva durante l'elaborazione della relativa incassata (che introduce il nome attrattore), in modo che quando il verbo principale appare la sua connessione con il soggetto corretto è già disponibile. Il nome attrattore non interferisce perché non è collegato alla stessa rete di oscillatori beta del soggetto principale: esso è connesso all'oscillatore theta che realizza la relativa incassata, che è a sua volta annidato all'interno del ciclo delta che realizza la frase principale. La gerarchia temporale separa strutturalmente i due livelli di elaborazione, impedendo l'interferenza del nome attrattore.

Nel Transformer, questo meccanismo non esiste: l'attenzione è globale e non strutturata gerarchicamente, e il nome attrattore può influenzare i pesi di attenzione che determinano l'accordo del verbo principale proprio per la sua maggiore prossimità lineare. Il sistema può imparare, statisticamente, che i nomi

attrattori non devono essere usati come target di accordo — e i modelli più grandi e meglio addestrati mostrano una riduzione sistematica di questo errore — ma lo fa attraverso la memorizzazione di pattern statistici, non attraverso l'implementazione di un vincolo strutturale. La differenza è visibile quando si testano strutture rarissime o inedite nel corpus di addestramento: il sistema biologico generalizza correttamente perché ha la struttura; il sistema statistico fallisce perché non ha i pattern.

Questa analisi non è una critica ai modelli linguistici in quanto strumenti ingegneristici: sono straordinariamente potenti per gli scopi pratici per cui vengono impiegati. È una precisazione teorica sulla natura dei processi che li realizzano e sulla loro rilevanza come modelli della cognizione linguistica umana. La proposta oscillatoria e i modelli Transformer non sono rivali: sono descrizioni di sistemi fondamentalmente diversi, ciascuno dei quali illumina, per contrasto, le proprietà dell'altro. E la specificità del sistema biologico — il fatto che usi il tempo come dimensione strutturale, che realizzi la gerarchia attraverso l'accoppiamento di frequenze, che integri il segnale linguistico con un modello incarnato del mondo — non è un difetto da correggere ma la chiave per comprendere perché il linguaggio umano sia quello che è.

XLII. Il giudizio di grammaticalità come fenomeno cosciente

Tra tutti i dati empirici di cui la linguistica teorica si avvale, il più fondamentale e al tempo stesso il più misterioso è il **giudizio di grammaticalità**: la capacità del parlante nativo di stabilire, con elevata concordanza tra parlanti e con elevata affidabilità, se una data stringa verbale è una frase della propria lingua o non lo è. Questo giudizio è immediato — si manifesta tipicamente senza deliberazione esplicita — ed è accompagnato da una sensazione qualitativa distintiva: la frase inaccettabile suona sbagliata, produce una certa resistenza fenomenologica, un senso di dissonanza che è difficile da descrivere ma facile da riconoscere.

Questa sensazione qualitativa — il *feeling* di grammaticalità o di inaccettabilità — è un fatto fenomenologico di primo ordine: è un'esperienza cosciente del parlante, non un'inferenza o una credenza teorica. Il parlante non *deduce* che la frase "*Il bambini ha mangiato la torta*" è agrammaticale applicando le regole dell'accordo morfologico: la *sente* agrammaticale, la sua agrammaticalità si impone alla coscienza con la stessa immediatezza con cui si impone il dolore o il suono di uno strumento stonato.

Questo fatto è al tempo stesso ovvio e profondamente problematico. Ovvio, perché i parlanti lo conoscono per esperienza diretta. Problematico, perché pone la domanda — che la proposta oscillatoria da sola non è in grado di rispondere — su come un processo computazionale, per quanto sofisticato, produca un'esperienza cosciente con le sue proprietà qualitative distintive. È il **problema**

difficile della coscienza di David Chalmers applicato al dominio specifico del linguaggio.

XLIII. La coscienza dell'oscillazione: Global Workspace e IIT

Le due teorie della coscienza più sviluppate e formalmente articolate nell'ambito delle neuroscienze cognitive — la teoria dello **spazio di lavoro globale** (*Global Workspace Theory*, GWT) di Bernard Baars e Stanislas Dehaene, e la teoria dell'**integrazione di informazione** (*Integrated Information Theory*, IIT) di Giulio Tononi — offrono strade diverse per connettere la proposta oscillatoria all'esperienza cosciente del linguaggio.

La GWT propone che la coscienza di un contenuto mentale corrisponda alla sua diffusione globale attraverso una rete di aree corticali ampiamente interconnesse — lo spazio di lavoro globale — che trasmette l'informazione localmente elaborata alle regioni specializzate di tutto il cervello. Questo accesso globale si distingue dall'elaborazione inconscia, che rimane confinata entro i moduli periferici. La firma neurofisiologica dell'accesso cosciente, nella GWT, è un'onda di attivazione corticale tardiva (circa 300–500 millisecondi dopo lo stimolo) caratterizzata da sincronizzazione beta-gamma su lunghe distanze corticali — un'ignizione (*ignition*) dell'attività corticale distribuita che accompagna il momento in cui un contenuto diventa cosciente.

Nella proposta oscillatoria del linguaggio, l'ignizione dello spazio di lavoro globale corrisponde al momento in cui la struttura sintattica locale (elaborata in theta-gamma nelle aree temporali) viene integrata con il modello globale del discorso (mantenuto in delta-beta nelle reti frontali-parietali). Il giudizio di grammaticalità come esperienza cosciente sarebbe il correlato soggettivo di questa integrazione globale: il momento in cui la struttura della frase entra nello spazio di lavoro globale e diventa disponibile per la deliberazione, la valutazione e il reporting.

Questo schema spiega perché il giudizio di grammaticalità sia immediato ma non istantaneo: i processi oscillatori locali che realizzano il parsing sintattico possono procedere inconsciamente e a grande velocità, mentre la coscienza del risultato richiede l'ignizione dello spazio di lavoro globale, che arriva con un ritardo di alcune centinaia di millisecondi. È questa latenza — ben documentata dai paradigmi ERP — che separa l'elaborazione subconscia della struttura sintattica dal giudizio cosciente di accettabilità.

L'IIT propone invece un approccio radicalmente diverso: la coscienza è identica alla quantità di informazione integrata dal sistema, misurata dalla grandezza Φ (phi). Un sistema ha un'esperienza cosciente con carattere qualitativo specifico se e soltanto se ha un valore Φ elevato — se, cioè, i suoi elementi sono fortemente

integrati in modo causalmente irriducibile a qualunque partizione del sistema in sottosistemi.

La definizione formale di Φ richiede di considerare tutte le possibili partizioni del sistema in due sottoinsiemi (A, B) e di misurare la quantità di informazione perduta da ciascuna partizione. Denotando con EI la *effective information* — la quantità di informazione trasmessa causalmente dal sistema — la partizione minima di informazione è:

$$\varphi^* = \min_{\{\text{partizioni}\}} EI(M \rightarrow M \mid M_0) - [EI(A \rightarrow A \mid A_0) + EI(B \rightarrow B \mid B_0)]$$

e il massimo su tutte le partizioni fornisce Φ , la misura dell'integrazione irriducibile. Un sistema con $\Phi = 0$ è completamente decomponibile e non ha esperienza cosciente; un sistema con Φ alto è irriducibilmente integrato e ha un'esperienza cosciente ricca.

Nel contesto della proposta oscillatoria, l'IIT suggerisce che l'esperienza cosciente del linguaggio — incluso il giudizio di grammaticalità come fenomeno qualitativo — corrisponda all'integrazione causalmente irriducibile delle oscillazioni a diverse scale temporali: il sistema delta-theta-beta-gamma è cosciente nella misura in cui la sua integrazione gerarchica è irriducibile alla somma delle sue parti. La sincronizzazione cross-frequenza non è soltanto un meccanismo computazionale: è il substrato dell'integrazione causale che genera la coscienza del senso linguistico.

XLIV. Il qualia della struttura: perché la frase agrammaticale suona sbagliata

C'è qualcosa di peculiare nell'esperienza qualitativa della frase agrammaticale che né la GWT né l'IIT, da sole, riescono a catturare completamente. Quando si incontra la frase "*Incolore idee verdi dormono furiosamente*" — il celebre esempio di Chomsky di frase grammaticalmente corretta ma semanticamente assurda — essa suona strana, ma di una stranezza diversa da quella di "*Il bambini ha mangiato la torta*". La prima viola aspettative semantiche ma rispetta la struttura sintattica; la seconda viola la struttura sintattica. Le due stranezze hanno un carattere qualitativo diverso — sono inconfondibili per qualunque parlante nativo — e questa differenza qualitativa è essa stessa un dato fenomenologico che qualunque teoria del linguaggio e della coscienza deve spiegare.

Nella prospettiva della codifica predittiva, le due stravaganze corrispondono a errori di previsione a livelli diversi della gerarchia: la frase di Chomsky produce un segnale di errore gamma al livello semantico (la co-occorrenza di *incolore* e *idee* con *verdi* e *dormono* viola le distribuzioni semantiche attese) senza produrre un

errore significativo al livello sintattico (la struttura soggetto-predicato è perfettamente regolare). La frase con accordo errato produce un segnale di errore al livello sintattico (l'accordo numero-persona è violato) senza necessariamente produrre un errore semantico (il contenuto proposizionale è perfettamente coerente). Il carattere qualitativo diverso delle due stranezze corrisponde al fatto che gli errori di previsione a livelli diversi della gerarchia hanno firme neurali diverse — diversi componenti ERP, diversi pattern oscillatori — e queste firme diverse si traducono in esperienze qualitative diverse.

Questo argomento suggerisce una proposta affascinante: i **qualia** — le proprietà qualitative dell'esperienza cosciente — non sono generici ma strutturati, e la loro struttura riflette la struttura della gerarchia predittiva del sistema. Il qualia del rosso e il qualia del blu differiscono perché corrispondono a errori di previsione a livelli diversi della gerarchia visiva; il qualia dell'agrammaticalità sintattica e il qualia dell'anomalia semantica differiscono perché corrispondono a errori a livelli diversi della gerarchia linguistica predittiva. La struttura dell'esperienza non è separata dalla struttura del sistema computazionale: ne è il risvolto soggettivo.

Questa proposta è radicale ma non arbitraria. È la versione, applicata al dominio linguistico, di una tesi più generale sulla relazione tra struttura dell'esperienza e struttura del sistema neurale che realizza l'esperienza — una tesi che risuona con alcune delle posizioni più sofisticate della filosofia della mente contemporanea, in particolare con il **naturalismo strutturale** di David Chalmers (la struttura dell'esperienza è isomorfa alla struttura funzionale del sistema) e con le proposte di Thomas Metzinger sull'esperienza come modellazione trasparente del sé (*Transparent Self-Model*).

XLV. L'ineffabilità del senso e i limiti della riduzione

Si è percorso un lungo cammino: dal formalismo di Merge all'equazione di Kuramoto, dalla plasticità hebbiana delle connessioni beta all'ignizione dello spazio di lavoro globale, dai qualia della struttura al naturalismo strutturale. Eppure, alla fine di questo percorso, rimane qualcosa che resiste — non la resistenza del mistero irrazionale, ma la resistenza di ciò che non si lascia catturare da nessun livello di descrizione senza residuo.

Questo residuo ha un nome nella tradizione filosofica: è l'**ineffabilità del senso**. Non il significato nel senso tecnico della semantica formale — il significato come condizione di verità, come funzione da mondi possibili a valori di verità — ma il senso nel senso più ampio e radicale di ciò che la parola apre per il parlante che la comprende davvero. La comprensione di una metafora, di un verso poetico, di una battuta ironica non si esaurisce nella sua descrizione funzionale: ha un surplus che la descrizione lascia fuori, un eccesso che è precisamente l'esperienza del senso come senso.

Wittgenstein, nelle *Ricerche filosofiche*, scriveva che comprendere una parola non è avere un'esperienza mentale che accompagna l'udirlo: è saper usarla, è padroneggiarla in tutti i contesti rilevanti, è sentirsi a casa nella rete di pratiche in cui essa si iscrive. Questa comprensione non è riducibile a nessuna singola esperienza né a nessun insieme finito di capacità funzionali: è una padronanza globale, olistica, che si manifesta nell'azione più che nell'introspezione.

La proposta oscillatoria, nella sua versione più ambiziosa, descrive le condizioni neurali che rendono possibile questa padronanza. Descrive il sistema che impara, che si sintonizza, che mantiene le dipendenze, che predice e corregge, che integra globalmente. Ma la padronanza stessa — il fatto che il parlante *sia* in casa nella lingua, non semplicemente *elabori* segnali linguistici — rimane un fatto di un ordine diverso da quello delle frequenze e degli accoppiamenti.

Questo non è uno scacco per la proposta oscillatoria: è il suo limite naturale, il confine che separa la scienza della cognizione linguistica dalla filosofia del linguaggio. Ogni teoria scientifica ha i propri presupposti e i propri limiti: non spiega la propria applicabilità, non giustifica i propri criteri di prova, non cattura la propria utilità. La proposta oscillatoria spiega molto — sorprendentemente molto, considerando quanto poco si comprendeva del substrato neurale della sintassi anche soltanto vent'anni fa. Ma non spiega perché il linguaggio *significhi*, perché la struttura apra un mondo condiviso tra parlanti, perché la parola giusta al momento giusto possa cambiare il corso di una vita.

Queste domande non sono fuori portata della filosofia: sono, da sempre, il suo terreno più proprio. E la proposta oscillatoria, nel delineare con precisione ciò che spiega e ciò che non spiega, ha il merito raro di restituire alla filosofia il suo dominio non per debolezza teorica ma per chiarezza epistemologica — la consapevolezza che le domande più profonde sul linguaggio non sono quelle che la scienza risolve, ma quelle che la scienza, nel suo progresso, rende più nitide e più urgenti.

XLVI. Ciò che abbiamo guadagnato

Conviene ora raccogliere i fili di questo lungo percorso e chiedersi, con serena franchezza, che cosa la proposta oscillatoria abbia effettivamente guadagnato rispetto allo stato del problema all'inizio. Non si tratta di stilare un consuntivo trionfalistico: si tratta di delineare con precisione il profilo di ciò che è stato argomentato, distinto da ciò che rimane speculativo o aperto.

Il guadagno principale è di natura concettuale, e risiede nella capacità della proposta di fornire un **livello di descrizione intermedio** tra la teoria formale della struttura sintattica e la biofisica dei neuroni. La sintassi chomskiana descriveva Merge come un'operazione set-teorica atemporale su oggetti astratti; la neuroscienza topografica descriveva le aree corticali del linguaggio in termini

di flusso sanguigno e lesioni. Tra questi due livelli vi era un vuoto computazionale — nessuna teoria di *come* un'operazione astratta fosse implementata da un sistema fisico temporalmente esteso.

La proposta oscillatoria riempie questo vuoto collocando la struttura sintattica nell'asse temporale delle oscillazioni corticali, trasformando la gerarchia formale in gerarchia dinamica e la computazione set-teorica in sincronizzazione di fase. Non è ancora una bridging theory nel senso pieno del termine — non riduce la sintassi alle oscillazioni né le oscillazioni alla biofisica — ma è una teoria dei *vincoli di implementazione*: stabilisce quali proprietà il substrato fisico deve avere per realizzare le operazioni formali, e quali proprietà formali sono determinate, e non soltanto permesse, dall'architettura oscillatoria.

Questo livello di descrizione intermedio è esattamente il tipo di teoria che le neuroscienze cognitive richiedono per progredire: non una riduzione eliminativista che dissolve la struttura nell'implementazione, né un dualismo che tiene separati i due livelli senza spiegare la loro relazione. È, per usare il termine di John Marr, il livello **algoritmico** della spiegazione: il livello che descrive *come* il compito computazionale viene eseguito, a metà strada tra la descrizione del compito (livello computazionale: la sintassi genera strutture gerarchiche) e la descrizione dell'implementazione (livello implementazionale: i neuroni scaricano con certi pattern).

XLVII. I debiti teorici aperti

Con uguale franchezza, occorre riconoscere i debiti teorici che la proposta ha contratto senza ancora saldare. Essi sono di due tipi: debiti empirici, che riguardano risultati sperimentali che la proposta predice ma che non sono ancora stati ottenuti; e debiti teorici, che riguardano aspetti della struttura formale che la proposta non riesce ancora a trattare in modo soddisfacente.

Tra i debiti empirici, il più urgente riguarda la **specificità sintattica** del segnale oscillatorio, già menzionata nella parte precedente. Distinguere la firma oscillatoria della struttura sintattica da quella della struttura prosodica, ritmica e semantica — che condividono molte delle stesse scale temporali — richiede paradigmi sperimentali molto più sofisticati di quelli attualmente disponibili. Un paradigma ideale dovrebbe presentare frasi in cui struttura sintattica e struttura prosodica siano sistematicamente dissociate — frasi con struttura sintattica complessa ma prosodia semplice, e viceversa — e misurare le differenze nei pattern di accoppiamento cross-frequenza nelle due condizioni. Tecnicamente, questo è possibile; ma la produzione di stimoli adeguati e il controllo delle variabili confondenti sono sfide non banali.

Un secondo debito empirico riguarda i **dati longitudinali sullo sviluppo**. Se la proposta acquisitiva è corretta, l'architettura oscillatoria del bambino dovrebbe

mostrare cambiamenti sistematici correlati con le tappe dell'acquisizione grammaticale: l'insorgenza dell'accoppiamento delta-theta dovrebbe coincidere con la comparsa delle prime strutture frasali; l'accoppiamento beta a lungo raggio dovrebbe consolidarsi nel periodo di acquisizione delle dipendenze a lungo raggio. Questi dati richiedono studi longitudinali EEG su popolazioni ampie di bambini nelle varie fasi dell'acquisizione — studi metodologicamente difficili ma non impossibili, e la cui realizzazione sistematica è ancora in larga misura davanti a noi.

Tra i debiti teorici, il più profondo riguarda il problema della **composizionalità semantica** in chiave oscillatoria. La proposta fin qui sviluppata si è concentrata sulla struttura sintattica — le relazioni di dipendenza tra costituenti — senza affrontare il problema di come il significato composizionale emerga dall'operazione di Merge. Merge combina non soltanto forme sintattiche ma anche contenuti semantici: il significato di "*un uomo alto*" è composizionalmente derivato dal significato di "*un uomo*" e di "*alto*" attraverso l'operazione di predicazione. Come questa composizione semantica viene realizzata nell'architettura oscillatoria? La proposta della codifica predittiva gerarchica fornisce un primo abbozzo di risposta, ma il dettaglio della composizione — la derivazione del significato dell'intero dal significato delle parti attraverso le regole semantiche — rimane da formalizzare in termini oscillatori.

XLVIII. Il programma di ricerca: tre fronti aperti

Alla luce dei guadagni e dei debiti, si può delineare il programma di ricerca che la proposta oscillatoria articola in modo più perspicuo di quanto qualunque alternativa attualmente disponibile sia in grado di fare. Questo programma si dispiega su tre fronti principali, ciascuno dei quali richiede la collaborazione di tradizioni disciplinari tipicamente separate.

Il **primo fronte** è sperimentale e neurofisiologico. Si tratta di sviluppare paradigmi per misurare, con la risoluzione spazio-temporale necessaria, le firme oscillatorie della struttura sintattica — distinguendole dalle firme prosodiche, ritmiche e semantiche. Questo richiede l'integrazione di EEG ad alta densità, MEG, e laddove possibile ECoG su pazienti neurochirurgici, con paradigmi linguistici accuratamente costruiti sulla base delle previsioni della proposta. Richiede anche lo sviluppo di metodi di analisi più sofisticati: oltre al PAC e al PLV, strumenti come la **Mutual Information** tra bande di frequenza, la **Transfer Entropy** per misurare la direzionalità del flusso di informazione, e le tecniche di analisi di rete sulle misure di connettività oscillatoria.

Il **secondo fronte** è computazionale e formale. Si tratta di costruire modelli computazionali espliciti — non metaforici — che implementino operazioni sintattiche come Merge attraverso dinamiche di sincronizzazione di oscillatori, e che producano previsioni quantitative sui pattern di PAC e PLV attesi per diverse

strutture sintattiche. Questi modelli devono essere falsificabili: devono produrre previsioni che, se disattese, portino alla revisione del modello. Il formalismo del modello di Kuramoto gerarchico con ritardi, abbozzato in questo saggio, è un punto di partenza, ma richiede estensioni significative per diventare biologicamente realistico e linguisticamente adeguato.

Il **terzo fronte** è filosofico e teorico. Si tratta di chiarire le implicazioni concettuali della proposta per le grandi questioni della filosofia del linguaggio e della mente: il problema dell'ontologia della struttura sintattica (è reale a livello neurale o è un'astrazione teorica?), il problema del grounding semantico (come le oscillazioni si connettono al significato?), il problema della coscienza linguistica (come la dinamica oscillatoria produce l'esperienza cosciente del senso?). Queste domande non sono risolvibili sperimentalmente: richiedono lavoro concettuale raffinato che prenda sul serio sia i risultati empirici che le tradizioni filosofiche rilevanti — la fenomenologia husserliana, la filosofia del linguaggio ordinario di Wittgenstein e Austin, la semantica formale di Montague e Kaplan, la filosofia della mente di Fodor, Chalmers e Metzinger.

XLIX. Una nota metodologica finale: l'unità della conoscenza e i rischi della sintesi

Un saggio come questo — che tenta di connettere la teoria formale della sintassi, la biofisica delle oscillazioni neurali, la matematica dei sistemi dinamici, la fenomenologia husserliana e la filosofia del linguaggio — corre un rischio metodologico che è utile riconoscere esplicitamente. Il rischio è quello di produrre una **sintesi prematura**: una visione d'insieme apparentemente coerente che nasconde, sotto la fluidità della narrazione, lacune argomentative, salti inferenziali e generalizzazioni indebite.

Questo rischio non è eliminabile: qualunque proposta teorica ambiziosa lo affronta. Ma può essere mitigato attraverso una pratica di onestà intellettuale che distingua con precisione tra ciò che è stabilito empiricamente, ciò che è plausibile e supportato da evidenza indiretta, ciò che è speculativo ma falsificabile, e ciò che è puramente ipotetico. Nel corso di questo saggio, si è cercato di applicare questa distinzione costantemente — pur riconoscendo che, nella scrittura saggistica come in quella scientifica, la linea tra speculazione illuminante e affermazione indebita è spesso difficile da tracciare con precisione.

Ciò che si è sostenuto con maggiore fiducia è il seguente nucleo: la struttura gerarchica della sintassi pone un problema di implementazione neurale che le neuroscienze topografiche non possono risolvere; le oscillazioni corticali forniscono un substrato temporale che in linea di principio permette di codificare gerarchia strutturale attraverso gerarchia di frequenze; l'accoppiamento cross-frequenza, misurato attraverso PAC e PLV, è sistematicamente modulato dalla struttura sintattica dell'input linguistico; il formalismo del modello di

Kuramoto gerarchico fornisce un quadro matematico in cui questo accoppiamento può essere analizzato in termini di sincronizzazione, transizioni di fase e attrattori dinamici.

Ciò che rimane più speculativo — e che viene presentato come tale — è la corrispondenza precisa tra livelli della gerarchia oscillatoria e livelli della gerarchia sintattica; il ruolo specifico delle oscillazioni beta nel mantenimento delle dipendenze a lungo raggio da Merge interno; l'identità tra la struttura degli attrattori del sistema dinamico e la competenza grammaticale del parlante; la relazione tra la gerarchia oscillatoria e i qualia dell'esperienza linguistica.

La proposta oscillatoria, nella sua interezza, non è ancora una teoria ma un **programma di ricerca** — nel senso di Lakatos: un nucleo duro di ipotesi non falsificabili per convenzione metodologica, circondato da un'euristica positiva che guida lo sviluppo di teorie particolari, e da un'euristica negativa che protegge il nucleo dalle falsificazioni dirette orientando le revisioni verso la cintura protettiva esterna. Il nucleo duro è l'ipotesi che la struttura sintattica sia implementata da struttura temporale oscillatoria; l'euristica positiva è lo sviluppo di modelli computazionali e paradigmi sperimentali che esplicitino questa ipotesi; l'euristica negativa è la rinuncia a falsificazioni basate su singole misure EEG in paradigmi non adeguatamente controllati.

Un programma di ricerca è fruttuoso se genera nuovi problemi e nuove soluzioni più velocemente di quanto ne accumuli di irrisolti. La proposta oscillatoria, giudicata in questi termini, appare promettente: i problemi che genera — sulla specificità sintattica del segnale oscillatorio, sulla composizionalità semantica, sullo sviluppo longitudinale, sulla variazione tipologica — sono problemi empiricamente accessibili e teoricamente fertili. Questo non è una garanzia di successo, ma è il miglior indicatore disponibile di una proposta teorica che vale la pena perseguire.

L. Chiusura: il cervello che narra

Vi è, alla fine di ogni riflessione profonda sul linguaggio, un momento in cui il pensiero si volge su se stesso e riconosce la propria condizione. Questo saggio è stato scritto in linguaggio; è stato prodotto da un sistema che elabora struttura linguistica; pretende di spiegare, almeno parzialmente, la natura del processo che lo ha generato. Vi è in questo qualcosa di vertiginoso e di irriducibilmente paradossale — non nel senso logico di una contraddizione formale, ma nel senso fenomenologico di una ricorsività esperienziale: il linguaggio che parla di se stesso attraverso se stesso.

Il cervello che oscilla, che si sincronizza, che minimizza la propria sorpresa, che mantiene le dipendenze attraverso il ritmo del beta — questo cervello è anche il cervello che narra. Che costruisce storie, che attribuisce intenzioni, che interpreta

il mondo attraverso trame di causa ed effetto, che si racconta una storia su se stesso che chiama "identità". Il linguaggio non è soltanto lo strumento della narrazione: è la condizione di possibilità del narratore, il medium attraverso il quale un sistema biologico diventa soggetto di un'esperienza che include il proprio passato e il proprio futuro come dimensioni costitutive del presente.

La proposta oscillatoria, per quanto elaborata, non raggiunge questo livello. Spiega il parsing, la segmentazione, l'integrazione, il mantenimento delle dipendenze. Non spiega perché il cervello che fa tutto questo sia anche un cervello che *si racconta* ciò che fa, che ne fa un'esperienza, che la custodisce nella memoria come parte di una vita che ha senso. Questo residuo — il senso della vita che il linguaggio rende possibile — è, forse, il problema più antico e più moderno di tutta la filosofia: non una domanda a cui rispondere, ma la forma stessa dell'interrogazione umana, che ogni risposta parziale rende più chiara senza mai esaurire.

Il cervello oscilla. Le oscillazioni si accoppiano in gerarchie. Le gerarchie realizzano strutture. Le strutture portano significati. I significati costruiscono mondi. I mondi sono abitati da esseri che si domandano come tutto questo sia possibile.

Questa catena non si chiude. Ed è precisamente per questo che il linguaggio continuerà ad essere, per chiunque voglia capire davvero che cosa siamo, il luogo più difficile e più necessario del pensiero.

LI. Il principio di composizionalità e il suo imbarazzo neurale

Il principio di composizionalità — che il significato di un'espressione complessa è determinato in modo sistematico dal significato dei suoi costituenti e dalla struttura sintattica che li combina — è uno dei pilastri più stabili della semantica formale, e al tempo stesso uno dei problemi più resistenti della neurolinguistica computazionale. Stabile, perché senza composizionalità non si spiegherebbe come i parlanti comprendano frasi mai udite prima: un bambino di sei anni che sente per la prima volta la frase "Il drago color crema ha distrutto il castello di ghiaccio" la comprende immediatamente, e questa comprensione non può essere spiegata senza ammettere che egli sappia combinare i significati dei costituenti secondo la struttura sintattica. Resistente, perché tradurre questo principio in un meccanismo neurale concreto richiede di spiegare non soltanto come il cervello rappresenti i significati lessicali, ma come li *componga* in tempo reale seguendo la struttura gerarchica della frase che si dispiega nell'input.

La difficoltà specifica è la seguente. I significati lessicali sono codificati in reti semantiche distribuite — tipicamente nelle corteccie temporali e parietali — dove le rappresentazioni di concetti simili sono vicine e quelle di concetti distanti sono lontane, in uno spazio semantico continuo la cui geometria riflette le

co-occorrenze distribuzionali nel linguaggio naturale e le similarità percettivo-motorie degli oggetti. Questa codifica distribuita è biologicamente plausibile e sperimentalmente documentata. Ma la composizione sintattica richiede di combinare questi vettori distribuiti secondo regole che *dipendono dalla struttura* — non dalla semplice somma o media dei vettori componenti, che produrrebbe la stessa rappresentazione per "il cane morde l'uomo" e "l'uomo morde il cane". Come fa il cervello a tenere traccia di *quale* costituente svolge *quale* ruolo nella struttura sintattica, mantenendo distinte le rappresentazioni di frasi con gli stessi elementi lessicali in relazioni strutturali diverse?

LII. Binding e sincronizzazione: la soluzione oscillatoria al problema della composizione

La proposta oscillatoria al problema della composizione semantica ha radici nel concetto di **binding** — la capacità del sistema nervoso di legare tra loro rappresentazioni distribuite che appartengono allo stesso oggetto o evento. Il problema del binding, nelle sue formulazioni più classiche nel dominio visivo, riguarda il fatto che diverse proprietà di un oggetto — colore, forma, movimento, posizione — sono elaborate in aree corticali diverse e devono essere integrate in un'unica rappresentazione coerente. La soluzione proposta da Wolf Singer e collaboratori — e ampiamente dibattuta, senza essere mai definitivamente risolta — è che il binding sia realizzato dalla **sincronizzazione gamma** tra le popolazioni neurali che elaborano le diverse proprietà: le rappresentazioni che appartengono allo stesso oggetto oscillano in fase, mentre quelle che appartengono a oggetti diversi oscillano fuori fase.

Nel dominio linguistico, il problema del binding è strutturalmente più complesso perché le rappresentazioni da legare non sono soltanto quelle di un oggetto percettivo, ma quelle dei ruoli tematici — chi fa cosa a chi — in una struttura predicativa. Nella frase "il cane morde l'uomo", il significato di *cane* deve essere legato al ruolo di Agente del predicato *mordere*, e il significato di *uomo* deve essere legato al ruolo di Paziente dello stesso predicato. Nella frase "l'uomo morde il cane", gli stessi significati lessicali devono essere legati ai ruoli inversi. La differenza tra le due rappresentazioni composizionali non è una differenza nel contenuto dei vettori lessicali — *cane* e *uomo* sono identici nelle due frasi — ma una differenza nel modo in cui questi vettori sono legati alla struttura predicativa del verbo.

La proposta oscillatoria per questo problema è la seguente. Le oscillazioni gamma che codificano il significato lessicale dei costituenti si sincronizzano selettivamente con le oscillazioni theta che codificano la posizione strutturale di quel costituente nell'albero sintattico. Il costituente in posizione di Agente si sincronizza con la fase theta corrispondente al ciclo di elaborazione del soggetto; il costituente in posizione di Paziente si sincronizza con la fase theta corrispondente al ciclo di elaborazione dell'oggetto. Il *binding*

semantico-strutturale è realizzato dalla coincidenza di fase: i neuroni che codificano il significato di *cane* scaricano in sincronia gamma nella stessa fase theta in cui viene elaborata la posizione di Agente.

Formalmente, si può descrivere questo meccanismo introducendo un operatore di binding fase-selettivo. Sia φ_A la fase theta associata alla posizione strutturale di Agente e φ_P quella associata alla posizione di Paziente. Sia $S_\alpha(t)$ il segnale gamma che codifica il significato del costituente α . Il binding di α al ruolo di Agente corrisponde alla condizione:

$$\langle S_\alpha(t) \cdot \cos(\theta_{\text{theta}}(t) - \varphi_A) \rangle_t > \varepsilon$$

dove $\theta_{\text{theta}}(t)$ è la fase istantanea dell'oscillazione theta e ε è una soglia di sincronizzazione. Questa condizione è soddisfatta quando il segnale gamma del costituente α è sistematicamente più forte nelle fasi theta vicine a φ_A — ossia quando l'ampiezza gamma di α è massima nella finestra temporale theta associata alla posizione di Agente. Il binding è dunque una forma specializzata di PAC theta-gamma: non una semplice modulazione dell'ampiezza gamma da parte della fase theta, ma una modulazione *fase-selettiva*, in cui diversi costituenti sono legati a diverse fasi dello stesso ciclo theta.

Questa proposta ha una conseguenza misurabile: frasi in cui i ruoli tematici non seguono il prototipo canonico agente-prima-paziente — come le frasi passive, le frasi con oggetto dislocato a sinistra, o le frasi con Experiencer come soggetto ("Mario teme i cani") — dovrebbero produrre pattern di PAC theta-gamma diversi rispetto alle frasi canoniche, perché il binding tra significato lessicale e posizione strutturale deve essere stabilito per via non prototipica, richiedendo un riallineamento delle fasi theta rispetto alle posizioni strutturali attese. Questa previsione è in linea con evidenze ERP che mostrano componenti N400 e P600 maggiori per frasi non canoniche, ma non è stata ancora testata direttamente attraverso misure di PAC.

LIII. La struttura dei ruoli tematici come geometria dello spazio di fase

Un modo alternativo e più geometrico di concepire la composizione semantica oscillatoria è attraverso il concetto di **spazio di fase** delle oscillazioni theta. Il ciclo theta — la finestra temporale di circa 125–250 millisecondi in cui un'unità sintattica viene elaborata — non è un intervallo indifferenziato: è diviso internamente in fasi distinte, ciascuna associata a un tipo di informazione o a un ruolo strutturale specifico.

Questa idea è ispirata dalla ricerca di György Buzsáki e collaboratori sull'ippocampo, dove i cicli theta sono suddivisi in finestre gamma successive, ciascuna delle quali codifica una diversa rappresentazione nell'ambito di una

sequenza ordinata — il cosiddetto **theta-gamma code**. In questa architettura, la sequenza temporale delle finestre gamma all'interno di un ciclo theta crea un codice di posizione: la rappresentazione codificata nella prima finestra gamma occupa la prima posizione, quella nella seconda finestra gamma occupa la seconda posizione, e così via. La posizione nella sequenza è codificata dalla fase theta, non dall'identità dei neuroni che scaricano.

Esteso al dominio sintattico, questo schema suggerisce che i ruoli tematici — Agente, Paziente, Tema, Destinatario, Strumento — siano codificati come **fasi specifiche** del ciclo theta, e che il significato lessicale di ciascun costituente venga associato al ruolo corrispondente attraverso la sincronizzazione gamma nella fase theta appropriata. La composizione semantica è, in questo senso, una scrittura nell'agenda spaziotemporale del ciclo theta: la struttura dei ruoli tematici è la geometria dello spazio di fase.

Questa concezione ha implicazioni interessanti per il problema della **saturazione argomentale** — il fatto che un verbo come *mordere* richiede necessariamente due argomenti (un Agente e un Paziente) e non può essere usato con uno solo senza produrre una frase incompleta. La saturazione argomentale corrisponde, in termini oscillatori, al riempimento di tutte le fasi theta previste dalla struttura del predicato: il verbo *mordere* apre due finestre di fase — una per l'Agente e una per il Paziente — e la frase è completa soltanto quando entrambe le finestre sono state riempite da un costituente appropriato. La sensazione di incompletezza che si prova udendo "il cane morde" in assenza di un complemento oggetto è il correlato fenomenologico di una finestra di fase aperta che attende il proprio contenuto — una protenzione husserliana realizzata come fase theta vuota in attesa di saturazione gamma.

LIV. Composizione come trasformazione del paesaggio semantico

Vi è un terzo modo di concepire la composizione semantica in chiave oscillatoria, complementare ai due precedenti e forse teoricamente più profondo: come trasformazione del paesaggio degli attrattori nello spazio delle rappresentazioni semantiche.

Nelle reti semantiche distribuite, i significati lessicali corrispondono ad attrattori — stati stabili della rete verso cui il sistema converge in risposta a input lessicali specifici. Quando il sistema elabora la parola *cane*, converge verso l'attrattore *cane* nello spazio delle rappresentazioni semantiche: una configurazione di attività neurale distribuita che cattura le proprietà prototipiche del concetto. Quando elabora la parola *morde*, converge verso l'attrattore *mordere*, che include non soltanto le proprietà motorie e fisiche del mordere ma anche la struttura di ruoli tematici del verbo.

La composizione semantica, in questa prospettiva, è la trasformazione che la struttura sintattica opera sul paesaggio degli attrattori: il sintagma verbale "*morde l'uomo*" non è la sovrapposizione degli attrattori di *mordere* e *uomo*, ma un nuovo attrattore — la rappresentazione della scena di mordere che ha *l'uomo* come Paziente — che emerge dalla modulazione reciproca tra i due attrattori lessicali mediata dalla struttura sintattica. Questa modulazione è realizzata dall'accoppiamento theta-gamma: la fase theta che segnala la posizione di Paziente modifica l'attrattore del verbo, selezionando la configurazione compatibile con un Paziente animato e umano.

Formalmente, questo può essere espresso come una trasformazione dipendente dalla struttura dell'attrattore predicativo. Se A_v è l'attrattore del verbo e A_α è l'attrattore del costituente α , la rappresentazione compositiva del sintagma verbale non è $A_v + A_\alpha$ ma:

$$A_{\{VP\}} = T_{\{\varphi_P\}}(A_v, A_\alpha)$$

dove $T_{\{\varphi_P\}}$ è un operatore di trasformazione che dipende dalla fase theta φ_P della posizione di Paziente — un operatore che seleziona, dalla varietà degli stati possibili dell'attrattore del verbo, quello compatibile con il contenuto del Paziente α . Questo operatore non è commutativo: $T_{\{\varphi_A\}}(A_v, A_\alpha) \neq T_{\{\varphi_P\}}(A_v, A_\alpha)$, il che riflette la non-commutatività semantica che differenzia "il cane morde l'uomo" da "l'uomo morde il cane". La non-commutatività della composizione semantica — un fatto semantico elementare, ma cruciale — emerge come conseguenza della geometria non commutativa dello spazio di fase theta.

Questo risultato è, a mio giudizio, uno dei punti più significativi dell'intera proposta: mostra che la distinzione tra strutture semanticamente diverse con gli stessi costituenti non è una proprietà ad hoc aggiunta al sistema, ma una conseguenza necessaria della geometria delle oscillazioni attraverso cui la composizione è realizzata. La struttura sintattica non è dunque soltanto un algoritmo che guida l'assemblaggio dei significati: è la trasformazione geometrica che crea nuovi significati compositivi che non erano presenti negli attrattori lessicali di partenza.

LV. La metafora come composizione non letterale: una finestra aperta

Il caso più suggestivo e teoricamente più sfidante per qualunque teoria della composizione semantica è quello della **metafora**. Quando si dice "Giulia è il sole della mia vita", il significato non è compositivo nel senso letterale: il valore di verità della frase non dipende dall'estensione dei predicati *essere* e *sole* applicati all'argomento *Giulia*. Eppure la frase ha un significato preciso — e un significato che emerge dalla struttura sintattica e dal contenuto lessicale, non è arbitrario. La

metafora è una forma di composizione non letterale che sfrutta la struttura per creare un significato nuovo che non era disponibile in nessuno dei costituenti presi separatamente.

Nella prospettiva degli attrattori, la metafora corrisponde a una situazione in cui la trasformazione $T_{\{\varphi\}}$ degli attrattori lessicali produce un attrattore compositazionale che non è vicino a nessuno degli attrattori lessicali di partenza, ma si colloca in una regione dello spazio semantico raggiungibile soltanto attraverso la composizione. L'attrattore della metafora "Giulia è il sole della mia vita" non è l'attrattore di *Giulia* né quello di *sole* né la loro combinazione letterale: è una regione nuova dello spazio semantico — irradiazione affettiva, centralità esistenziale, vitalità — che la struttura predicativa di identità porta in esistenza *ex novo* attraverso la composizione degli attrattori lessicali disponibili.

Questo schema ha una previsione empirica interessante: la comprensione della metafora dovrebbe richiedere un periodo di elaborazione più lungo rispetto alla comprensione letterale, perché il sistema deve trovare un attrattore in una regione del paesaggio semantico non immediatamente accessibile dagli attrattori di partenza — deve, in un senso preciso, *creare* un nuovo attrattore piuttosto che convergere verso uno preesistente. Le evidenze ERP sulla metafora — che mostrano N400 più ampie per metafore nuove rispetto a metafore convenzionalizzate, con profilo temporale compatibile con una ricerca di attrattore prolungata — sono coerenti con questa previsione, anche se il collegamento con le dinamiche oscillatorie specifiche rimane da esplorare direttamente.

La composizione metaforica è, in definitiva, l'esempio più puro del fatto che il significato non è contenuto nei costituenti: emerge dalla struttura che li combina, e può emergere in forme che nessuno dei costituenti poteva anticipare. Questo è il motivo per cui il linguaggio è inesauribile come strumento di creazione concettuale — e perché qualunque teoria che riduca il significato a proprietà preesistenti dei lessemi, senza spiegare come la struttura generi proprietà nuove, è fondamentalmente inadeguata. La proposta oscillatoria, nel trattare la composizione come trasformazione geometrica del paesaggio degli attrattori, fornisce una cornice in cui questo aspetto creativo del linguaggio non è un'eccezione da gestire ma una conseguenza necessaria dell'architettura del sistema.

LVI. Il problema del pensiero verbale e il suo imbarazzo strutturale

Vi è una forma di esperienza linguistica che qualunque teoria del linguaggio deve confrontare, e che risulta tra le più difficili da collocare nel quadro neuroscientifico: il **linguaggio interno**, il discorso silenzioso che si svolge nella mente senza produzione vocale. Quando si rilegge mentalmente una frase, quando si pianifica un argomento prima di esporlo, quando ci si trova a "sentire"

una voce nella testa durante la lettura silenziosa — questi fenomeni costituiscono il linguaggio nel suo modo di esistenza più intimo e meno esposto all'osservazione empirica.

La difficoltà teorica è duplice. Sul versante fenomenologico: in che misura il linguaggio interno è genuinamente linguistico, ossia strutturato sintatticamente, piuttosto che una forma di rappresentazione concettuale fluida che ha soltanto l'apparenza del linguaggio? Sul versante neurale: se il linguaggio interno non implica produzione vocale — e dunque nessun segnale acustico esterno con cui sincronizzarsi — come si realizzano le oscillazioni corticali che, nella proposta oscillatoria, sono normalmente entrained al ritmo del parlato?

Queste due difficoltà sono in realtà aspetti dello stesso problema, e la proposta oscillatoria offre una via per affrontarli congiuntamente. Ma per percorrerla occorre prima chiarire il quadro teorico in cui il linguaggio interno si iscrive — e questo richiede di confrontarsi con due tradizioni che hanno affrontato il problema da angolature molto diverse: la tradizione vygotskijana dello sviluppo del linguaggio interno, e la tradizione fodoriana del mentalese.

LVII. Vygotskij e l'internalizzazione della parola

Lev Semënovič Vygotskij, nei suoi scritti della metà degli anni Trenta — raccolti postumo in *Pensiero e linguaggio* (1934) — aveva proposto una genesi sviluppa del linguaggio interno che rappresenta ancora oggi una delle intuizioni più profonde e meno esaurite della psicologia cognitiva. Per Vygotskij, il linguaggio interno non è una semplice riduzione del linguaggio esterno — il parlato interiore come il parlato esteriore ma senza voce — ma un sistema semiotico radicalmente diverso nelle sue proprietà strutturali, prodotto da un processo di **internalizzazione** progressiva del dialogo con l'altro.

Il bambino piccolo, nella ricostruzione vygotskijana, pensa ad alta voce — il suo linguaggio è esterno e socializzato. Progressivamente, con lo sviluppo, il linguaggio si contrae e si interiorizza: prima appare il linguaggio egocentrico (il bambino parla tra sé), poi il linguaggio interno vero e proprio, in cui la struttura sintattica esterna si riduce a qualcosa di molto più condensato, predicativo e orientato alla prospettiva del soggetto. Il linguaggio interno vygotskijano non è frase compiuta: è frammento, tensione, predicato senza soggetto esplicito, perché il soggetto è sempre implicito nell'orizzonte esperienziale del pensante.

Questa descrizione fenomenologica — il linguaggio interno come struttura ridotta, predicativa, implicitamente prospettica — ha un correlato nella proposta oscillatoria che vale la pena esplicitare. Se il linguaggio esterno richiede la sincronizzazione delle oscillazioni corticali con il segnale acustico esterno, e se questa sincronizzazione si svolge dalla periferia verso il centro (bottom-up), il linguaggio interno funzionerebbe in modo invertito: le oscillazioni corticali non

sono entrained da un segnale acustico esterno ma generate autonomamente dalle regioni corticali superiori e proiettate verso il basso (top-down) attraverso le connessioni discendenti della gerarchia predittiva. Il linguaggio interno è, in questa prospettiva, una forma di **previsione senza correzione**: il sistema genera il proprio input linguistico attraverso le connessioni discendenti, senza il contrappeso delle correzioni ascendenti che il segnale acustico reale fornirebbe.

La struttura contratta e predicativa del linguaggio interno vygotskijano si spiegherebbe allora come conseguenza dell'assenza del segnale di errore esterno: in assenza di un input acustico con cui riconciliarsi, il sistema non ha bisogno di mantenere espliciti tutti i livelli della gerarchia sintattica — i livelli inferiori (fonologico, sillabico) sono soppressi o ridotti a un livello minimale di attivazione, e restano attivi principalmente i livelli superiori (sintagmatico, frasale) che sono necessari per la pianificazione e la valutazione del contenuto. Il linguaggio interno è strutturalmente ridotto perché soltanto una parte della gerarchia oscillatoria è necessariamente impegnata.

LVIII. Il mentalese di Fodor e la critica oscillatoria

La tradizione fodoriana affronta il problema del linguaggio interno da una prospettiva diametralmente opposta. Jerry Fodor, nel *The Language of Thought* (1975), aveva proposto che il pensiero avvenga in un linguaggio mentale — il **mentalese** — che è strutturalmente simile al linguaggio naturale (ha simboli, struttura compositiva, ricorsività) ma è distinto da esso: è il medium del pensiero, non il medium della comunicazione. Il linguaggio naturale, in questa prospettiva, è semplicemente la traduzione del mentalese in un codice comunicabile — un'interfaccia verso l'esterno, non il pensiero stesso.

La proposta fodoriana ha il merito di spiegare perché il pensiero sia strutturato e compositivo senza identificarlo con il linguaggio naturale specifico che un parlante parla. Ma paga un prezzo teorico alto: postula l'esistenza di un sistema simbolico non identificato con nessun substrato fisico noto, con proprietà computazionali specifiche che non possono essere spiegate in termini di meccanismi neurali conosciuti. Il mentalese è una risorsa esplicativa che moltiplica le entità — aggiunge un livello computazionale al di sopra delle rappresentazioni neurali senza specificare come le due si relazionino — in violazione del rasoio di Occam metodologico.

La proposta oscillatoria offre una risposta alla domanda fodoriana che non richiede di postulare un sistema simbolico separato. Il medium del pensiero verbale non è il mentalese, né il linguaggio naturale nel suo pieno dispiegamento fonologico, ma la **gerarchia oscillatoria corticale** nella sua modalità endogena — la stessa architettura di accoppiamento cross-frequenza che realizza il parsing del linguaggio esterno, operante però in modalità top-down, senza il vincolo dell'input acustico esterno.

In questa prospettiva, pensare in italiano e pensare in giapponese non è la stessa cosa — contro l'intuizione di Fodor, che sosteneva l'indipendenza del mentale dalla lingua naturale acquisita — perché l'architettura oscillatoria è specificatamente sintonizzata sulla lingua acquisita: il parlante italiano pensa con oscillatori beta che codificano le dipendenze strutturali dell'italiano, con un paesaggio di attrattori plasmato dalle strutture sintattico-semantiche dell'italiano. Non nel senso grossolano e insostenibile della relatività linguistica radicale alla Whorf — che i concetti disponibili dipendano dalla lingua — ma nel senso più sottile e empiricamente accessibile che la *geometria* del pensiero verbale interno dipende dall'architettura oscillatoria, e questa architettura è parzialmente determinata dalla lingua acquisita.

LIX. La rete di default e il substrato neurale del monologo interiore

Le neuroscienze cognitive hanno identificato negli ultimi vent'anni una rete corticale — la **Default Mode Network** (DMN) — che si attiva durante il riposo, la divagazione mentale, il pensiero autobiografico, la pianificazione del futuro e la comprensione narrativa, e che si deattiva durante compiti cognitivi che richiedono attenzione focalizzata verso l'esterno. La DMN include il cortex prefrontale mediale, la corteccia cingolata posteriore, il precuneo, e la giunzione temporoparietale — regioni che si sovrappongono parzialmente ma non completamente con le reti del linguaggio.

Il rapporto tra la DMN e il linguaggio interno è diventato oggetto di interesse crescente. Studi di fMRI mostrano che durante il pensiero verbale spontaneo — la divagazione mentale che include frequentemente monologo interiore, pianificazione verbale, rievocazione narrativa — la DMN è coattivata con le regioni del linguaggio nelle corteccie temporali e prefrontali inferiori. Ma la natura di questa coattivazione — se si tratti di elaborazione genuinamente sintattica o di attivazione lessicale senza struttura — è ancora oggetto di dibattito.

La proposta oscillatoria aggiunge un livello di descrizione che il fMRI, con la sua bassa risoluzione temporale, non può catturare: la domanda cruciale non è *dove* si attiva la DMN durante il pensiero verbale, ma *come* le sue oscillazioni si accoppiano con quelle delle reti del linguaggio. Se il monologo interiore è genuinamente strutturato sintatticamente, ci si aspetta che durante la divagazione verbale le oscillazioni della DMN mostrino accoppiamento con le oscillazioni theta-beta delle reti linguistiche — accoppiamento che avviene in modalità top-down, con le regioni prefrontali della DMN che guidano le oscillazioni theta delle corteccie temporali attraverso le connessioni discendenti, in assenza di un input acustico che guidi il ritmo.

Studi di MEG sul riposo e sulla divagazione mentale mostrano che la DMN è caratterizzata da oscillazioni prominenti nelle bande alpha e beta —

precisamente le bande associate alle predizioni discendenti nella proposta di Bastos. Questo è compatibile con l'idea che la DMN sia il substrato del pensiero top-down — il sistema che genera rappresentazioni interne in assenza di input esterno — e che il suo coinvolgimento nel linguaggio interno corrisponda alla generazione endogena della gerarchia oscillatoria linguistica, dalla sua struttura discorsiva (delta infra-lento della DMN) fino ai livelli sintagmatici (beta della rete prefrontale-temporale) e lessicali (theta-gamma delle corteccie temporali).

LX. Patologia del linguaggio interno: il caso delle voci udite

Un banco di prova particolarmente istruttivo per la proposta oscillatoria del linguaggio interno è quello delle **allucinazioni uditive verbali** — le "voci" che molti pazienti con schizofrenia, e occasionalmente soggetti non psicotici in condizioni di stress estremo o di privazione sensoriale, riferiscono di udire come provenienti dall'esterno. Queste voci hanno proprietà linguistiche precise: hanno un'identità, un'intonazione, talvolta un accento; producono frasi complete e semanticamente coerenti; spesso commentano le azioni del paziente o gli danno ordini. Eppure nessuno parla: il segnale acustico esterno è assente.

L'interpretazione più accreditata delle allucinazioni uditive verbali le concepisce come forme di linguaggio interno che il paziente non riconosce come autoprodotte — un failure del *source monitoring*, la capacità di distinguere i propri pensieri dalla percezione esterna. In termini di codifica predittiva, le allucinazioni corrisponderebbero a predizioni top-down sufficientemente forti da "invadere" il livello percettivo senza essere bilanciate dai segnali di errore ascendenti — un estremo della stessa distorsione del rapporto predizione-correzione che, in forma più lieve, caratterizza la psicosi.

In termini oscillatori, questa proposta si traduce in una previsione specifica: durante le allucinazioni uditive verbali, le regioni dell'area di Broca e delle corteccie temporali superiori dovrebbero mostrare oscillazioni top-down (beta) senza il corrispondente segnale di errore ascendente (gamma), producendo un pattern oscillatorio che assomiglia all'elaborazione del linguaggio esterno ma senza la componente bottom-up che normalmente accompagna l'input acustico reale. Alcune evidenze EEG e MEG su pazienti durante gli episodi allucinatori sono compatibili con questa previsione — in particolare la documentazione di oscillazioni gamma nelle regioni temporali superiori *senza* il corrispondente segnale acustico esterno che normalmente le genera — ma i dati rimangono ancora frammentari e metodologicamente difficili da raccogliere.

Il caso delle allucinazioni uditive verbali illustra un aspetto della proposta oscillatoria che vale la pena sottolineare esplicitamente: il confine tra linguaggio esterno e linguaggio interno non è assoluto ma corrisponde a un continuum di rapporti tra le componenti bottom-up e top-down della gerarchia oscillatoria. Il linguaggio esterno è dominato dal bottom-up (il segnale acustico trascina le

oscillazioni corticali); il linguaggio interno normale è dominato dal top-down (le corteccie superiori generano oscillazioni senza segnale esterno) ma include ancora segnali propriocettivi dall'apparato articolatorio (le cosiddette *subvocalizzazioni* — i leggeri movimenti muscolari della laringe durante il pensiero verbale silenzioso, documentati dalla EMG della laringe). Le allucinazioni verbali si collocano in un punto estremo del continuum: top-down puro, senza né input acustico esterno né propriocezione articolatoria, eppure percepito come esterno.

Questa collocazione sul continuum non è soltanto una descrizione: è una previsione di come il sistema debba funzionare in condizioni intermedie — meditazione, trance, stati ipnagogici, stati di flusso — che la proposta oscillatoria rende empiricamente accessibili attraverso le misure di PAC e di direzionalità del flusso oscillatorio. Il rapporto tra pensiero e linguaggio, in questa prospettiva, è un rapporto tra modalità diverse di equilibrio tra le componenti ascendenti e discendenti della gerarchia oscillatoria — non una distinzione discreta tra due sistemi separati, ma un parametro continuo che la fisiologia delle oscillazioni rende misurabile.

LXI. Il problema del parallelo musicale

La musica e il linguaggio condividono una caratteristica che li distingue nettamente dalla maggior parte degli altri sistemi cognitivi: entrambi sono sequenze temporali di elementi discreti dotati di struttura gerarchica, entrambi sono acquisiti per esposizione durante un periodo critico dello sviluppo, entrambi attivano reti corticali ampiamente sovrapposte nell'emisfero sinistro e destro, entrambi producono risposte elettrofisiologiche caratteristiche alla violazione delle aspettative strutturali. Eppure musica e linguaggio sono chiaramente distinti: la musica non ha contenuto proposizionale, non denota stati di cose nel mondo, non possiede ruoli tematici né accordo morfologico né dipendenze di controllo. Un accordo di settima di dominante non ha il suo *referente* nel mondo fisico nel modo in cui la parola *cane* denota cani.

Questa somiglianza strutturale con differenza funzionale fa della musica il banco di prova ideale per isolare ciò che è specifico della proposta oscillatoria applicata al linguaggio da ciò che è invece proprietà generale di qualunque sistema gerarchico temporalmente esteso elaborato dal cervello umano. Se l'architettura oscillatoria della gerarchia delta-theta-beta-gamma è specifica al linguaggio, ci si aspetta che la musica recluti una gerarchia oscillatoria strutturalmente diversa. Se invece è una proprietà generale del sistema cognitivo per l'elaborazione di strutture gerarchiche temporali, ci si aspetta una gerarchia simile — con le stesse bande di frequenza, lo stesso schema di accoppiamento cross-frequenza — adattata alle proprietà specifiche della struttura musicale.

La risposta che i dati offrono è né l'una né l'altra nella forma netta: vi è una gerarchia oscillatoria nell'elaborazione musicale che è strutturalmente analoga a

quella linguistica per alcune proprietà e che si discosta da essa per altre, in modo sistematico e teoricamente istruttivo. Esaminare queste analogie e divergenze è l'obiettivo di questa parte.

LXII. La gerarchia metrica come accoppiamento delta-theta in musica

La struttura ritmica della musica è organizzata in modo gerarchico: i battiti (beats) sono raggruppati in misure (bars), le misure in frasi musicali, le frasi in periodi, i periodi in sezioni formali. Questa gerarchia ritmica è analoga — almeno strutturalmente — alla gerarchia prosodica del linguaggio: i battiti musicali corrispondono alle sillabe accentate, le misure alle parole o ai piedi metrici, le frasi musicali alle frasi prosodiche. Non è sorprendente, dunque, che la gerarchia oscillatoria dell'elaborazione musicale rispecchi questa struttura metrica: le oscillazioni corticali si sincronizzano con la struttura ritmica della musica a diverse scale temporali, con le oscillazioni theta che si agganciano al ritmo dei singoli battiti e le oscillazioni delta che si agganciano al ritmo delle misure e delle frasi.

Questo schema è stato documentato con notevole precisione da Nozaradan e collaboratori, in una serie di studi di EEG che hanno utilizzato stimoli ritmici di complessità variabile per analizzare selettivamente i diversi livelli della gerarchia metrica. Presentando ai partecipanti sequenze ritmiche in complessità 4/4 — il metro quaternario della musica popolare occidentale — le oscillazioni corticali mostrano picchi di potenza a multipli della frequenza fondamentale del battito: al battito (theta, ≈ 2 Hz per un tempo di 120 bpm), alla semimisura (≈ 4 Hz), alla misura (≈ 1 Hz), e alla sopramisura (≈ 0.5 Hz). Crucialmente, l'ampiezza relativa di questi picchi dipende non soltanto dalla struttura acustica dello stimolo — il ritmo fisico dei suoni — ma dalla struttura metrica *percepita*: lo stesso stimolo acustico può evocare metre diverse (3/4 vs 6/8), e le oscillazioni corticali rispecchiano la metrica percepita, non soltanto quella acustica.

Questo risultato ha due implicazioni importanti. La prima è che l'accoppiamento delta-theta nell'elaborazione musicale non è semplicemente un riflesso passivo della struttura acustica dello stimolo: è il prodotto di un processo generativo attivo, in cui le previsioni del sistema interno sul ritmo atteso modulano la risposta al ritmo acustico — esattamente come nella proposta linguistica le previsioni della gerarchia sintattica modulano la risposta al segnale acustico del parlato. La seconda implicazione è che la metrica musicale, come la struttura prosodica del linguaggio, è rappresentata nel cervello come una struttura di accoppiamento oscillatorio gerarchico — non come una serie di note o di tempi, ma come una rete di connessioni di fase tra oscillatori a frequenze diverse.

LXIII. La sintassi musicale: ERAN, P600 e la violazione tonale

La musica tonale occidentale ha una forma di struttura che analogicamente si chiama **sintassi tonale**: le note non si susseguono in modo arbitrario ma secondo progressioni armoniche vincolate dalla tonalità della composizione. In una composizione in do maggiore, certi accordi sono attesi (l'accordo di dominante sol maggiore che risolve verso do maggiore), altri sono inattesi ma ammissibili (la sostituzione di tritono, il napolitano), altri ancora sono fortemente dissonanti con la tonalità stabilita. Questa struttura di aspettative armoniche, acquisita implicitamente dai parlanti della cultura musicale occidentale attraverso l'esposizione, produce risposte elettrofisiologiche alle violazioni analoghe — ma non identiche — a quelle prodotte dalle violazioni sintattiche nel linguaggio.

La componente ERP più caratteristica delle violazioni armoniche in musica è la **ERAN** (*Early Right Anterior Negativity*): una negatività che emerge circa 150–200 millisecondi dopo un accordo armonicamente incongruente, localizzata nelle regioni frontali anteriori destre, seguito spesso da una seconda componente negativa più tardiva (ELAN musicale). Questa componente è sensibile alla distanza armonica dell'accordo target dalla tonalità stabilita: accordi più lontani producono ERAN più ampie, in modo sistematico e graded.

La ERAN ha caratteristiche che la avvicinano alla ELAN sintattica linguistica (*early left anterior negativity*), che si osserva circa 100–200 millisecondi dopo una violazione di categoria sintattica nel linguaggio. Entrambe le componenti riflettono l'elaborazione rapida di violazioni strutturali prima dell'integrazione semantica. Ma le differenze sono rivelatrici: la ERAN è prevalentemente destra, la ELAN è prevalentemente sinistra; la ERAN riflette struttura armonica (relazioni tra altezze), la ELAN riflette struttura sintattica (relazioni tra categorie grammaticali). Le due violazioni reclutano reti corticali parzialmente diverse, con asimmetrie emisferiche diverse — compatibili con la proposta di Poeppel che l'emisfero sinistro sia specializzato per l'elaborazione a grana fine (segmentale, morfologica) e l'emisfero destro per l'elaborazione a grana grossa (prosodica, armonica).

Nelle fasi più tardive dell'elaborazione, le violazioni armoniche in musica producono una componente P600 — apparentemente identica a quella prodotta dalle violazioni sintattiche nel linguaggio. Questo ha alimentato l'ipotesi di un **Shared Syntactic Integration Resource** (SSIR) tra musica e linguaggio — una risorsa computazionale condivisa per l'integrazione di strutture gerarchiche, indipendente dal dominio specifico (tonale o sintattico). L'ipotesi SSIR, formulata da Stefan Koelsch e da Stefan Frisch, prevede interferenza tra elaborazione musicale e linguistica quando le due avvengono simultaneamente: se le risorse di integrazione sono condivise, elaborare una violazione armonica mentre si elabora una frase sintattica complessa dovrebbe aumentare il tempo di reazione o ridurre la qualità del parsing di entrambe.

Le evidenze sperimentali sull'interferenza musica-linguaggio sono miste: alcuni studi trovano interferenza, altri non la trovano, e la direzione e l'entità dell'effetto dipendono dalla complessità degli stimoli e dalle condizioni sperimentali. In

termini oscillatori, l'ipotesi SSIR si tradurrebbe nell'esistenza di oscillatori o connessioni condivisi tra le reti della musica e quelle del linguaggio — un'ipotesi che potrebbe essere testata misurando se le perturbazioni indotte in beta (il livello dell'integrazione strutturale) da stimoli musicali interferiscono con il PAC sintattico durante l'elaborazione linguistica simultanea.

LXIV. La melodia come struttura temporale senza referenza

La differenza più profonda tra musica e linguaggio, dal punto di vista della proposta oscillatoria, non è nella struttura della gerarchia di accoppiamento — che è analoga — ma nella presenza o assenza di ciò che si potrebbe chiamare **ancoramento referenziale** della struttura. Le oscillazioni che realizzano la gerarchia sintattica linguistica sono, attraverso la composizione semantica e il binding theta-gamma, ancorate a contenuti concettuali che si riferiscono a stati di cose nel mondo. Le oscillazioni che realizzano la gerarchia metrica e tonica della musica non hanno questo ancoraggio: il PAC theta-gamma che elabora la struttura ritmica non lega le proprie finestre di fase a rappresentazioni di oggetti, agenti o eventi; le predizioni discendenti della sintassi tonale non sono predizioni su stati del mondo ma su stati dello spazio armonico.

Questa differenza non è triviale. Implica che la musica realizzi la struttura senza significato referenziale — o, più precisamente, con un tipo di significato che è interamente immanente alla struttura stessa, senza uscire verso il mondo. L'accordo di dominante "vuole" risolvere verso la tonica, non nel senso in cui il sintagma nominale soggetto "vuole" un verbo per completare la frase — una relazione di dipendenza che si riferisce agli agenti e agli eventi del mondo — ma nel senso in cui una tensione nello spazio armonico genera un'aspettativa di rilascio che è interna alla struttura tonale.

Questo è il motivo per cui la musica può esprimere emozioni ma non proposizioni: le sue strutture oscillatorie realizzano geometrie di tensione e rilascio, aspettativa e sorpresa, nell'asse temporale — ma queste geometrie non denotano. L'emozione musicale sarebbe, in questa prospettiva, il correlato soggettivo di certi pattern di errore di previsione nella gerarchia oscillatoria tonale: la tensione armonica come previsione frustrata, la risoluzione come previsione soddisfatta, la modulazione armonica inattesa come errore di previsione tonale. Queste dinamiche hanno un valore affettivo che non dipende dalla referenza — e questo è precisamente ciò che rende la musica capace di toccare emozioni in assenza di contenuto proposizionale.

Eduard Hanslick, nel suo trattato *Vom Musikalisch-Schönen* (1854), aveva sostenuto che la musica non esprime emozioni specifiche ma soltanto il movimento dinamico dell'emozione — la sua forma temporale, il suo profilo di intensificazione e rilascio — indipendentemente dal suo contenuto specifico. Questa proposta, formulata in termini estetici ottocenteschi, trova un'inaspettata

risonanza nella proposta oscillatoria: la musica elabora la struttura dinamica del tempo senza ancoraggio referenziale, e questa struttura dinamica pura è il substrato delle risposte affettive che la musica produce. La proposta oscillatoria fornisce, per la prima volta, un quadro neurale entro cui questa intuizione di Hanslick può essere tradotta in previsioni empiriche testabili.

LXV. Il contrasto rivelatore: cosa la musica ci dice del linguaggio

L'analisi comparativa musica-linguaggio, nella prospettiva oscillatoria, illumina per contrasto due aspetti del linguaggio che, trattando il linguaggio da solo, rischiano di passare inosservati.

Il primo aspetto è la **specificità del binding semantico-strutturale**. Il fatto che le oscillazioni theta-gamma del linguaggio realizzino un binding tra posizioni strutturali (fasi theta) e contenuti concettuali referenziali (popolazioni gamma semantiche) — mentre la musica utilizza lo stesso schema di accoppiamento senza binding referenziale — rivela che il binding è una proprietà aggiuntiva che il linguaggio costruisce *sopra* l'architettura oscillatoria di base. Il linguaggio non è soltanto una struttura temporale gerarchica: è una struttura temporale gerarchica *con ancoraggio referenziale*. Questo ancoraggio richiede, presumibilmente, connessioni oscillatorie specifiche tra le reti temporali del linguaggio e le reti semantiche concettuali — connessioni che la musica non sviluppa allo stesso modo, o le sviluppa in modo diverso (verso rappresentazioni affettive e motorie piuttosto che concettuali).

Il secondo aspetto è la **necessità della ricorsività**. La musica ha struttura gerarchica ma non ricorsività in senso chomskyan: non esiste una composizione musicale in cui una frase musicale sia contenuta all'interno di se stessa — la ricorsività che permette alla sintassi linguistica di produrre strutture di profondità arbitraria. La musica può avere gerarchie profonde — sezioni di forma sonata che contengono esposizioni che contengono temi che contengono frasi che contengono motivi — ma queste gerarchie non sono generate da una singola operazione ricorsiva come Merge: sono costruite attraverso convenzioni formali che variano tra stili e periodi storici.

In termini oscillatori, questa differenza si traduce nella domanda se la musica recluti il terzo livello della gerarchia oscillatoria — la modulazione del theta da parte del delta — nello stesso modo in cui il linguaggio lo fa per realizzare la struttura frasale. Studi di accoppiamento cross-frequenza durante l'ascolto musicale mostrano modulazioni delta-theta simili a quelle linguistiche per la struttura metrica delle frasi musicali, ma con una caratteristica importante: nella musica, il ciclo delta che corrisponde alla frase musicale è *periodico e prevedibile* — la struttura metrica della maggior parte della musica tonale occidentale è regolare, con frasi di 4 o 8 battiti ricorrenti in modo quasi meccanico. Nella sintassi linguistica, il ciclo delta che corrisponde alla frase è *aperiodico* — le frasi hanno

durate variabili e non predicibili dalla sola struttura metrica, perché la loro durata dipende dalla struttura sintattica della frase specifica.

Questa differenza — periodicità prevedibile in musica vs aperiodica in linguaggio — è la firma oscillatoria della ricorsività. La ricorsività produce strutture di lunghezza arbitraria e non prevedibile; la non-ricorsività produce strutture regolari e metricamente prevedibili. Il delta oscillatorio si sincronizza con la struttura frasale in entrambi i casi, ma lo fa in modalità diverse: entrained periodicamente dalla struttura metrica in musica, aggiornato aperiodicamente dalla struttura sintattica in linguaggio. Questa differenza nelle modalità di sincronizzazione delta è la traduzione oscillatoria della differenza computazionale tra grammatiche regolari (musica) e grammatiche libere da contesto (linguaggio) — una distinzione che, nella gerarchia chomskyana delle grammatiche formali, separa i sistemi finiti a stati dai sistemi ricorsivi.

Il contrasto con la musica rivela dunque, per via negativa, ciò che è essenziale al linguaggio e non si riduce alle proprietà generali dell'elaborazione gerarchica temporale: il binding referenziale e la ricorsività sono le proprietà che distinguono la gerarchia oscillatoria del linguaggio dalla gerarchia oscillatoria della musica — e sono, non casualmente, le proprietà che la biolinguistica chomskyana ha identificato come costitutive della Facoltà del Linguaggio nel suo senso stretto. La proposta oscillatoria non contraddice la teoria chomskyana: la traduce in un linguaggio neurofisiologicamente implementabile, specificando i meccanismi mediante i quali le proprietà formali astratte si realizzano nell'architettura temporale del cervello oscillante.

LXVI. Le lingue dei segni: struttura senza suono

Il silenzio più eloquente dell'intero saggio è stato quello sulle **lingue dei segni**. La loro assenza non è stata occasionale: è il riflesso di un imbarazzo teorico che occorre affrontare esplicitamente, perché le lingue dei segni costituiscono il banco di prova più radicale immaginabile per la proposta oscillatoria. Se la gerarchia sintattica è realizzata dall'accoppiamento tra oscillazioni corticali entrained al segnale acustico del parlato — dalla sincronia sillabica theta fino alla sincronia frasale delta — allora cosa avviene nel cervello del sordo segnante, che non ha mai udito il parlato e che elabora il linguaggio attraverso la modalità visuo-gestuale? Non vi è segnale acustico con cui sincronizzarsi; non vi è ritmo sillabico da agganciare alla banda theta; non vi è inviluppo temporale dell'onda sonora da seguire con le oscillazioni delta. Eppure le lingue dei segni sono lingue complete — hanno fonologia, morfologia, sintassi, pragmatica — e i sordi segnanti nativi mostrano le stesse asimmetrie emisferiche, le stesse aree corticali attivate, la stessa sensibilità alle violazioni strutturali che si osservano nei parlanti udenti.

Come si concilia questa evidenza con la proposta oscillatoria? La risposta richiede di distinguere con cura tra ciò che nella proposta è essenziale e ciò che è contingente all'implementazione acustica.

Ciò che è essenziale è la **gerarchia temporale di accoppiamento cross-frequenza** come meccanismo per realizzare struttura gerarchica distribuita nel tempo. Ciò che è contingente è la sorgente del segnale che guida questa gerarchia: nella modalità acustica, è l'involuppo temporale del parlato; nella modalità visuo-gestuale, potrebbe essere una struttura temporale equivalente — il ritmo dei movimenti manuali, la struttura temporale delle configurazioni della mano, la prosodia gestuale che organizza il flusso del segno in unità prosodiche analoghe alle frasi intonative.

LXVII. La prosodia gestuale e l'entrainment visivo

Le lingue dei segni hanno una struttura prosodica propria, articolata in unità gerarchicamente organizzate: il **segno** come unità fonologica minima, l'**unità ritmica** (corrispondente alla parola o al morfema), il **gruppo di respirazione** (corrispondente alla frase prosodica), la **frase** (corrispondente alla clausola). Questa struttura prosodica è realizzata da parametri gestuali — la velocità del movimento, la tensione muscolare, le pause, l'ampiezza dell'escursione manuale — che variano sistematicamente ai confini delle unità prosodiche, in modo analogo alle variazioni di F0 e durata che segnalano i confini prosodici nel parlato.

La domanda cruciale, in termini oscillatori, è se il sistema corticale del segnante si sincronizzi con questa struttura prosodica gestuale attraverso le stesse bande di frequenza che il sistema del parlante si sincronizza con la struttura prosodica acustica. La risposta disponibile — in un numero ancora piccolo ma teoricamente significativo di studi di EEG e MEG su sordi segnanti nativi — suggerisce che la risposta sia sostanzialmente affermativa: le oscillazioni delta si sincronizzano con i confini delle frasi nei segni, e le oscillazioni theta mostrano una modulazione correlata con il ritmo delle unità gestuali. L'entrainment non avviene attraverso la via acustica ma attraverso la via visiva e propriocettiva — il sistema visivo che elabora i movimenti dell'interlocutore segnante, e il sistema propriocettivo che monitora i movimenti delle proprie mani durante la produzione.

Questo risultato ha un'implicazione di portata teorica considerevole: la gerarchia oscillatoria del linguaggio non è intrinsecamente acustica. È una proprietà del sistema corticale che elabora struttura linguistica gerarchica, indipendentemente dalla modalità — acustica o visuo-gestuale — attraverso cui quella struttura viene trasmessa. Il segnale acustico nel parlato e il segnale visuo-propriocettivo nella lingua dei segni sono due implementazioni diverse della stessa funzione: fornire un'impalcatura temporale ritmica con cui le oscillazioni corticali possano

sincronizzarsi, realizzando la gerarchia strutturale del linguaggio nell'asse temporale.

Questa conclusione rafforza, invece di indebolire, la proposta oscillatoria nella sua versione più ambiziosa: mostra che la gerarchia oscillatoria è davvero una proprietà della facoltà linguistica in senso biologico — una proprietà che si manifesta indipendentemente dalla modalità, come ci si aspetta da qualcosa che appartiene al nucleo della dotazione biologica per il linguaggio.

LXVIII. La fonologia spaziale: simultaneità e gerarchia

Le lingue dei segni pongono un ulteriore problema teorico che non ha equivalente nel parlato: la **simultaneità**. Mentre il parlato è fondamentalmente lineare — i fonemi si succedono nel tempo uno dopo l'altro — i segni possono esprimere simultaneamente più morfemi attraverso la configurazione della mano, la posizione nello spazio, il movimento e l'orientamento del palmo. Un singolo segno può incorporare il radicale del verbo, l'accordo soggetto-oggetto, e la valenza dell'azione in un'unica configurazione visuo-gestuale simultanea, laddove il parlato esprimerebbe gli stessi morfemi in sequenza temporale.

Questa simultaneità pone un problema per la proposta oscillatoria, che concepisce la struttura come codificata nell'asse temporale delle oscillazioni: se morfemi distinti sono espressi simultaneamente, come vengono separati nella gerarchia oscillatoria? La risposta che si può avanzare — ancora largamente ipotetica — è che nelle lingue dei segni la distinzione tra morfemi simultanei sia realizzata non attraverso la dimensione temporale ma attraverso la **dimensione spaziale** delle oscillazioni corticali: morfemi espressi simultaneamente da diverse parti del corpo o in diverse posizioni dello spazio peripersonale vengono elaborati da popolazioni neurali diverse, che oscillano in gamma a frequenze simili ma con relazioni di fase spazialmente distribuita.

Questa proposta introduce nella teoria oscillatoria una dimensione che fino a questo punto era rimasta sullo sfondo: la possibilità che non soltanto il tempo ma anche lo spazio corticale partecipi alla codifica della struttura. Le oscillazioni gamma sono generate da circuiti locali, e popolazioni diverse in regioni corticali diverse possono oscillare in gamma in modo relativamente indipendente. Nelle lingue dei segni, questa localizzazione spaziale delle oscillazioni gamma potrebbe essere il meccanismo attraverso cui la simultaneità morfologica viene gestita: anziché distribuire i morfemi nell'asse temporale (come nel parlato), li distribuisce nell'asse spaziale delle popolazioni oscillanti (come nella lingua dei segni). La struttura sintattica sovraordinata — la gerarchia delta-theta che organizza le frasi — rimane temporale anche nelle lingue dei segni, perché la struttura sintattica è fondamentalmente lineare anche in queste lingue, nonostante la morfologia simultanea.

Questa proposta — struttura sintattica nel dominio temporale delle oscillazioni, morfologia simultanea nel dominio spaziale delle oscillazioni gamma — è falsificabile: predice che le misure di PAC nelle lingue dei segni mostrino le stesse strutture temporali (delta-theta-beta) del parlato per la sintassi, ma differenze sistematiche nell'organizzazione spaziale delle oscillazioni gamma per la morfologia simultanea. È un programma di ricerca che non ha ancora ricevuto attenzione sistematica — e che potrebbe produrre risultati di grande importanza teorica.

LXIX. Il bilinguismo: due architetture oscillatorie in un cervello

Il parlante **bilingue** — colui che ha acquisito due lingue a livello di competenza nativa, o comunque elevato — offre una prospettiva sperimentale irripetibile sulla proposta oscillatoria. Se l'architettura oscillatoria è parzialmente plasmata dalla lingua acquisita, il bilingue che ha acquisito due lingue tipologicamente diverse dovrebbe avere due architetture oscillatorie parzialmente distinte — una per ciascuna lingua — che coesistono nello stesso cervello. La domanda è: come coesistono? Competono? Si sovrappongono? Sono attivate selettivamente in funzione della lingua in uso?

La risposta che la proposta oscillatoria suggerisce, e che le evidenze sperimentali cominciano a profilare, è la seguente: le architetture oscillatorie delle due lingue non sono completamente separate né completamente sovrapposte, ma si organizzano in un sistema a sovrapposizione parziale con selezione contestuale. Le componenti della gerarchia oscillatoria che corrispondono alle proprietà universali del linguaggio — la struttura delta-theta-beta-gamma come architettura fondamentale, l'accoppiamento fase-ampiezza come meccanismo di nesting gerarchico — sono condivise tra le due lingue. Le componenti che corrispondono alle proprietà specifiche di ciascuna lingua — il range preciso delle frequenze sintonizzate al ritmo sillabico, la geometria delle connessioni beta per le dipendenze strutturali caratteristiche di quella lingua, la presenza o assenza dello strato di coupling tonale — sono parzialmente distinte e selezionate contestualmente in funzione della lingua attiva.

La selezione contestuale è realizzata, con ogni probabilità, da un meccanismo di **gating neuromodulatorio**: quando il parlante bilingue passa da una lingua all'altra — il fenomeno del code-switching — il sistema noradrenergico o colinergico modifica globalmente il guadagno delle connessioni sinaptiche, favorendo il pattern di accoppiamento oscillatorio della lingua in uso e sopprimendo quello della lingua non in uso. Questo gating non è istantaneo: il code-switching richiede un intervallo di tempo misurabile — dell'ordine di decine o centinaia di millisecondi — per completarsi, e questo intervallo corrisponde, nella proposta oscillatoria, al tempo necessario perché il sistema oscillatorio si ristabilizzi nel nuovo attrattore — l'architettura oscillatoria della lingua verso cui si sta commutando.

LXX. Il code-switching come transizione di fase oscillatoria

Il **code-switching** — l'alternanza tra due lingue nel corso di una singola conversazione, a volte all'interno di una singola frase — è uno dei fenomeni più affascinanti e meno compresi del bilinguismo. Non è un errore: è una competenza linguistica sofisticata, soggetta a vincoli strutturali precisi (le lingue si alternano tipicamente ai confini di costituente, non al loro interno), e prodotta in modo fluido e automatico da bilingui competenti. La questione di come il cervello bilingue gestisca il code-switching senza perdere la struttura grammaticale di nessuna delle due lingue è teoricamente profonda.

Nella proposta oscillatoria, il code-switching corrisponde a una **transizione di fase** del sistema dinamico corticale: il passaggio da un attrattore oscillatorio (l'architettura della Lingua A) a un altro (l'architettura della Lingua B). Come tutte le transizioni di fase, questa transizione non è istantanea ma si svolge attraverso una regione critica — un breve intervallo in cui il sistema è lontano da entrambi gli attrattori, più sensibile alle fluttuazioni e più vulnerabile alle interferenze. Questo intervallo critico è il correlato dinamico del costo cognitivo del code-switching, documentato sperimentalmente come un aumento del tempo di reazione e una maggiore frequenza di errori nelle prove di denominazione immediatamente successive a un cambio di lingua.

La previsione oscillatoria è che questo intervallo critico sia associato a una riduzione temporanea del PAC nelle bande rilevanti — una "perdita di accoppiamento" mentre il sistema transisce tra i due attrattori — seguita dal ristabilirsi del PAC nella banda della lingua verso cui si sta commutando. Questa sequenza di perdita e ristabilimento del coupling dovrebbe essere visibile nell'EEG con una risoluzione temporale dell'ordine di decine di millisecondi — una previsione che non è ancora stata testata direttamente, ma che è metodologicamente accessibile con le tecniche disponibili.

Il code-switching intra-frasale — l'alternanza di lingua all'interno di una singola frase — è particolarmente istruttivo perché è soggetto a vincoli strutturali precisi. Il *Matrix Language Frame Model* di Carol Myers-Scotton propone che in ogni episodio di code-switching vi sia una lingua matrice che fornisce l'impalcatura grammaticale della frase (morfologia, ordine delle parole, struttura argomentale) e una lingua ospite i cui elementi lessicali vengono inseriti nelle posizioni strutturali previste dalla lingua matrice. In termini oscillatori, la lingua matrice corrisponderebbe all'architettura oscillatoria dominante nell'episodio di code-switching: le sue oscillazioni delta-theta-beta forniscono l'impalcatura temporale entro cui gli elementi lessicali della lingua ospite vengono inseriti, legandosi alle posizioni di fase theta appropriate attraverso il meccanismo di binding gamma. La lingua ospite contribuisce il contenuto lessicale (oscillazioni gamma nelle reti semantiche) senza alterare la struttura dell'impalcatura (oscillazioni delta-theta-beta della lingua matrice).

Questa proposta ha una previsione misurabile: durante il code-switching intra-frasale, il PAC delta-theta dovrebbe rispecchiare la lingua matrice anche quando gli elementi lessicali inseriti appartengono alla lingua ospite — il livello dell'impalcatura strutturale resta nell'architettura della lingua matrice, mentre il livello lessicale (gamma) commuta verso le rappresentazioni della lingua ospite. Testare questa previsione richiederebbe misure di EEG ad alta densità su bilingui durante la produzione spontanea di code-switching — un disegno sperimentale tecnicamente impegnativo ma non impossibile.

LXXI. L'età di acquisizione e la stratificazione oscillatoria

Una dimensione del bilinguismo che la proposta oscillatoria illumina in modo particolarmente elegante è quella dell'**età di acquisizione** della seconda lingua. È noto da decenni che i bilingui precoci — coloro che hanno acquisito entrambe le lingue durante il periodo critico, tipicamente entro i cinque o sei anni — mostrano profili neurolinguistici sistematicamente diversi dai bilingui tardivi: nei bilingui precoci, le due lingue attivano regioni corticali più sovrapposte e mostrano asimmetrie emisferiche simili a quelle dei monolingui; nei bilingui tardivi, la seconda lingua mostra un'attivazione corticale meno efficiente, più diffusa, con maggiore coinvolgimento delle regioni prefrontali e minore delle regioni temporali.

Nella prospettiva oscillatoria, questa differenza riflette la profondità e la separazione dei bacini di attrazione delle due architetture oscillatorie. Il bilingue precoce ha acquisito entrambe le lingue durante il periodo di massima plasticità oscillatoria — quando la costante di apprendimento $\lambda(t)$ era elevata — e ha dunque plasmato due architetture oscillatorie entrambe profondamente radicate nel paesaggio energetico corticale, con bacini di attrazione ben definiti e con una separazione netta che permette la selezione contestuale efficiente. Il bilingue tardivo ha acquisito la seconda lingua quando $\lambda(t)$ era già ridotta: l'architettura oscillatoria della seconda lingua si è sovrapposta su quella già consolidata della prima lingua, producendo un paesaggio energetico meno netto — i bacini di attrazione delle due lingue si sovrappongono parzialmente, la selezione contestuale è meno efficiente, e l'elaborazione della seconda lingua richiede un investimento computazionale maggiore per ristabilizzare il sistema nell'attrattore appropriato.

La previsione è che i bilingui tardivi mostrino PAC sistematicamente più basso nella seconda lingua rispetto ai bilingui precoci, anche a parità di competenza comportamentale — perché la profondità del bacino di attrazione non è catturata dalla performance superficiale ma soltanto dalle misure di dinamica oscillatoria. Questa previsione distingue la proposta oscillatoria da qualunque teoria basata soltanto sulla competenza: predice differenze nelle firme neurali anche in assenza di differenze comportamentali manifeste, in accordo con il

principio che la struttura degli attrattori non è riducibile alle traiettorie superficiali che essa determina.

LXXII. Il talamo come hub oscillatorio: un'assenza colpevole

Nel corso dell'intero saggio, il talamo è stato menzionato soltanto di sfuggita — come generatore delle oscillazioni lente nei circuiti cortico-talamici — senza ricevere la trattazione che la sua centralità fisiologica esigerebbe. Questa omissione è tanto più colpevole in quanto il talamo non è semplicemente un relé passivo che trasmette segnali dalla periferia alla corteccia: è un **hub oscillatorio attivo** che partecipa alla generazione, alla sincronizzazione e alla regolazione delle oscillazioni corticali a tutte le scale temporali rilevanti per il linguaggio. Ignorarlo equivale a descrivere l'architettura di una cattedrale ignorando i contrafforti: la struttura regge in superficie, ma manca del sostegno che la tiene in piedi.

Il talamo è organizzato in nuclei con connettività specifica verso diverse regioni corticali. I nuclei talamici di **proiezione specifica** — come il corpo genicolato mediale (MGN) per l'udito, il pulvinar per la visione, il nucleo ventrale anteriore e il nucleo ventrale laterale per i circuiti motori — inviano proiezioni topograficamente organizzate verso le cortecce sensoriali e motorie primarie. I nuclei di **proiezione aspecifica** — come il nucleo centromediano, il nucleo reticolare — inviano proiezioni diffuse verso ampie regioni corticali e svolgono un ruolo di regolazione globale del livello di eccitabilità corticale. Il **nucleo reticolare talamico** (nRT) è di particolare importanza: è un foglio di neuroni GABAergici che circonda il talamo e riceve collaterali dai neuroni talamo-corticali e cortico-talamici, inviando inibizione verso i nuclei talamici specifici. Il nRT è il generatore principale dei **sleep spindles** — le oscillazioni a 11–16 Hz che appaiono durante il sonno NREM — e partecipa alla regolazione delle oscillazioni alfa e beta durante la veglia.

Per il linguaggio, il nucleo talamico più direttamente rilevante è il **pulvinar**, il nucleo talamico di maggiori dimensioni nei primati, che riceve proiezioni da praticamente tutte le aree corticali associative e invia proiezioni verso le cortecce temporali superiori, parietali inferiori e prefrontali inferiori — le stesse regioni che costituiscono il nucleo della rete corticale del linguaggio. Il pulvinar non processa il segnale acustico primario — questo avviene nel MGN e nella corteccia uditiva primaria — ma svolge un ruolo di **gating attenzionale**: regola quali informazioni provenienti dalle cortecce sensoriali vengono amplificate e trasmesse verso le reti prefrontali per l'elaborazione di ordine superiore, e quali vengono sopprese.

In termini oscillatori, il pulvinar esercita questo gating attraverso la sincronizzazione selettiva: quando sincronizza le proprie oscillazioni con quelle di una particolare regione corticale, amplifica il segnale di quella regione

facilitando la trasmissione verso le aree prefrontali; quando non si sincronizza, sopprime il segnale. Durante l'elaborazione linguistica, si ipotizza che il pulvinar sincronizzi selettivamente le oscillazioni theta delle cortecce temporali sinistre — le regioni che elaborano la struttura sillabica e lessicale — facilitando la trasmissione di questa informazione verso le cortecce prefrontali inferiori che realizzano l'integrazione sintattica. Il gating pulvinare è dunque il meccanismo attraverso cui l'attenzione selettiva modula l'elaborazione linguistica: ascoltare in ambienti rumorosi, comprendere parlato emotivamente rilevante, elaborare frasi in condizioni di carico cognitivo elevato — tutte situazioni in cui la selezione attenzionale dell'informazione linguistica è cruciale — richiedono una modulazione del gating talamico che la proposta oscillatoria deve integrare.

LXXIII. Il circuito cortico-talamo-corticale: la matematica del loop

Il circuito cortico-talamo-corticale — la connessione bidirezionale tra corteccia e talamo che forma un loop chiuso — è il substrato anatomico delle oscillazioni a lungo raggio che organizzano le scale temporali delta e alfa/beta. La matematica di questo circuito merita di essere esplicitata, perché rivela come le frequenze di risonanza del loop — e dunque le bande di frequenza dominanti nelle oscillazioni corticali — siano determinate dai ritardi di conduzione e dalle proprietà biofisiche delle sinapsi coinvolte.

Si consideri un modello minimale del loop cortico-talamico: una popolazione corticale con attività media $x_C(t)$ e una popolazione talamica con attività media $x_T(t)$, connesse da proiezioni eccitatorie bidirezionali con ritardi τ_{CT} (cortico-talamico) e τ_{TC} (talamo-corticale), e con inibizione reticolare dal nRT con ritardo τ_{RT} . Le equazioni del moto linearizzate del sistema sono:

$$\begin{aligned}\tau_C \cdot \dot{x}_C(t) &= -x_C(t) + w_{TC} \cdot x_T(t - \tau_{TC}) \\ \tau_T \cdot \dot{x}_T(t) &= -x_T(t) + w_{CT} \cdot x_C(t - \tau_{CT}) - w_{RT} \cdot x_{nRT}(t - \tau_{RT}) \\ \tau_{nRT} \cdot \dot{x}_{nRT}(t) &= -x_{nRT}(t) + w_{C \rightarrow nRT} \cdot x_C(t - \tau_{CT}) + w_{T \rightarrow nRT} \cdot x_T(t)\end{aligned}$$

dove i pesi w_{ij} sono le forze di connessione sinaptiche e τ_i sono le costanti di tempo delle popolazioni. Per trovare le frequenze di oscillazione del sistema, si cerca la soluzione nella forma $x_C(t) = A_C \cdot e^{i\omega t}$, $x_T(t) = A_T \cdot e^{i\omega t}$, che sostituita nelle equazioni fornisce il polinomio caratteristico in ω :

$$\begin{aligned}D(\omega) &= (1 + i\omega\tau_C)(1 + i\omega\tau_T)(1 + i\omega\tau_{nRT}) - \\ &\quad w_{TC} \cdot w_{CT} \cdot e^{\{-i\omega(\tau_{TC} + \tau_{CT})\}} \cdot (1 + i\omega\tau_{nRT}) + \\ &\quad w_{RT} \cdot [w_{C \rightarrow nRT} \cdot e^{\{-i\omega\tau_{CT}\}} + w_{T \rightarrow nRT}] \cdot w_{TC} \cdot \\ &\quad e^{\{-i\omega\tau_{TC}\}} = 0\end{aligned}$$

Le radici complesse $\omega^* = \omega_r + i\omega_i$ di questo polinomio determinano le frequenze di oscillazione ($\omega_r/2\pi$) e i tassi di smorzamento (ω_i) del sistema. Per valori tipici dei parametri — $\tau_{\{CT\}} + \tau_{\{TC\}} \approx 20\text{--}40$ ms per i loop cortico-talamici a breve raggio, $\approx 80\text{--}120$ ms per i loop a lungo raggio — le radici del polinomio caratteristico cadono nelle bande alpha-beta (8–20 Hz) per i loop a breve raggio e nelle bande delta-theta (1–8 Hz) per i loop a lungo raggio.

Questo risultato — ricavato dall'analisi matematica delle proprietà del circuito, non postulato ad hoc — spiega perché le bande di frequenza dominanti delle oscillazioni corticali siano quelle osservate: sono le frequenze di risonanza dei loop cortico-talamici, determinate dalla geometria anatomica delle connessioni e dalle proprietà biofisiche delle sinapsi. Le bande delta-theta, che la proposta oscillatoria identifica come il substrato della gerarchia sintattica di frasi e di sintagma, corrispondono alle frequenze di risonanza dei loop cortico-talamici a lungo raggio — i loop che connettono le corteccie prefrontali e temporali attraverso il pulvinar. Le bande beta, associate al mantenimento delle dipendenze a lungo raggio, corrispondono alle risonanze dei loop a medio raggio tra corteccie frontali e parietali inferiori attraverso i nuclei talamici di proiezione specifici.

Il talamo non è dunque una struttura passiva nel sistema oscillatorio del linguaggio: è il risonatore che, attraverso la geometria delle sue connessioni e le proprietà delle sue sinapsi, determina le frequenze dominanti della gerarchia corticale. Modificare il talamo — attraverso lesioni, neuromodulazione farmacologica, o variazioni genetiche nelle sue proprietà biofisiche — modifica le frequenze di risonanza e dunque l'architettura oscillatoria che realizza la gerarchia sintattica. Le afasie talamiche — deficit linguistici prodotti da lesioni talamiche, documentati da decenni ma rimasti teoricamente oscuri — trovano in questa prospettiva una spiegazione coerente: non sono il risultato della distruzione di un "centro" del linguaggio talamico (il talamo non ha funzione lessicale o sintattica specifica), ma il risultato della perturbazione del gating oscillatorio che il talamo esercita sulle reti corticali del linguaggio, alterando le frequenze di risonanza dei loop cortico-talamici e dunque l'architettura della gerarchia oscillatoria.

LXXIV. La neuromodulazione: dopamina, acetilcolina e il guadagno oscillatorio

La proposta oscillatoria, così come è stata sviluppata, descrive l'architettura delle oscillazioni corticali come se essa fosse fissa — determinata dalla connettività anatomica e dalle proprietà biofisiche degli interneuroni. Ma le oscillazioni corticali sono in realtà dinamicamente regolate da sistemi di neuromodulazione che modificano le proprietà delle sinapsi e delle membrane neuronali su scale temporali di secondi o minuti, senza alterare la connettività anatomica di fondo. Ignorare questi sistemi è ignorare i meccanismi attraverso cui il cervello regola la propria architettura oscillatoria in funzione del contesto — attenzione, arousal,

emozione, apprendimento — e dunque in funzione delle esigenze computazionali dell'elaborazione linguistica in corso.

I sistemi neuromodulatori principali — dopaminergico, colinergico, noradrenergico, serotoninergico — esercitano effetti distinti sulle oscillazioni corticali, che possono essere formalizzati come modificazioni dei parametri del modello oscillatorio. La **dopamina**, rilasciata dai neuroni della substantia nigra e del nucleo tegmentale ventrale verso la corteccia prefrontale e il corpo striato, aumenta il guadagno delle connessioni sinaptiche eccitatorie nel circuito prefrontale, facilitando le oscillazioni gamma locali e il PAC theta-gamma nelle reti prefrontali. Il suo effetto sull'elaborazione sintattica è indiretto ma rilevante: aumentando il PAC theta-gamma nelle reti prefrontali, facilita il mantenimento delle dipendenze strutturali a lungo raggio che richiedono memoria di lavoro. La deplezione dopaminergica — come nel morbo di Parkinson — riduce sistematicamente il PAC prefrontale e compromette le capacità di memoria di lavoro, con conseguenze sull'elaborazione delle frasi con dipendenze a lungo raggio che sono compatibili con questa previsione.

L'**acetilcolina**, rilasciata dai neuroni del nucleo basale di Meynert verso la corteccia e dall'ippocampo verso il setto mediale e viceversa, esercita effetti complessi sulle oscillazioni corticali: sopprime le oscillazioni alfa nelle cortecce sensoriali (riducendo l'inibizione top-down e facilitando l'input sensoriale) e potenzia le oscillazioni gamma locali (aumentando l'eccitabilità degli interneuroni parvalbuminergici). Il suo effetto netto sull'elaborazione linguistica è di aumentare la precisione del segnale sensoriale ascendente — ridurre il rumore nelle componenti bottom-up della gerarchia predittiva — favorendo l'aggiornamento delle rappresentazioni sintattico-semantiche in risposta all'input acustico. Nelle prime fasi dell'acquisizione linguistica, quando il sistema deve essere massimamente sensibile all'input per plasmare l'architettura oscillatoria, i livelli di acetilcolina corticale sono particolarmente elevati — compatibilmente con l'idea che la neuromodulazione colinergica sia uno dei meccanismi che regolano la costante di plasticità $\lambda(t)$ discussa nella sezione sull'acquisizione.

La **noradrenalina**, rilasciata dal locus coeruleus verso praticamente tutta la corteccia, regola il livello globale di arousal e la selettività dell'attenzione. La sua curva dose-risposta ha la forma di un U invertito: a livelli bassi (sonno), sopprime tutte le oscillazioni; a livelli ottimali (attenzione focalizzata), aumenta il rapporto segnale-rumore amplificando il segnale degli oscillatori fortemente attivati e sopprimendo quello degli oscillatori debolmente attivati; a livelli eccessivi (stress acuto), produce un aumento caotico dell'eccitabilità che degrada la selettività. Il suo effetto sull'elaborazione linguistica è di modulare la profondità dei bacini di attrazione del paesaggio energetico grammaticale: a livelli ottimali, approfondisce i bacini delle strutture sintatticamente corrette e abbassa le barriere tra attrattori adiacenti, facilitando la convergenza rapida verso la struttura corretta durante il parsing. Nelle condizioni di stress o di alta arousal che degradano il processamento sintattico — difficoltà di comprensione in

situazioni emotivamente intense — questa curva dose-risposta offre una spiegazione fisiologica precisa del deterioramento osservato.

LXXV. Il sonno e il consolidamento della grammatica: la notte che plagia il giorno

Il territorio più eluso dell'intero saggio è probabilmente quello del **sonno** e del suo ruolo nel consolidamento della struttura grammaticale acquisita. Questo silenzio è ancora più sorprendente considerando che il sonno è il processo fisiologico in cui le oscillazioni corticali sono più vivide, più strutturate e più facilmente misurabili — e considerando che c'è una letteratura crescente, benché ancora relativamente giovane, sul consolidamento delle rappresentazioni linguistiche durante il sonno.

Il sonno NREM — e in particolare il **sonno a onde lente** (slow-wave sleep, SWS) — è caratterizzato da oscillazioni corticali di grande ampiezza nelle bande delta (0.5–2 Hz) e sub-delta (< 0.5 Hz), denominate **onde lente** (slow oscillations, SO), che riflettono alternanze ritmiche tra stati di depolarizzazione (stati ON, associati a scariche neuronali) e stati di iperpolarizzazione (stati OFF, in cui i neuroni sono silenziosi). Sovraimposti alle onde lente, durante la fase ON, appaiono i **sleep spindles** — oscillazioni a 11–16 Hz generate dal nucleo reticolare talamico — e le **sharp-wave ripples** ippocampali — brevi oscillazioni ad alta frequenza (80–100 Hz) associate alla riattivazione delle memorie recenti nell'ippocampo.

Questo triplice nesting — onde lente che organizzano spindles che organizzano ripples — è ora riconosciuto come il meccanismo principale del **consolidamento mnemonico** durante il sonno: le esperienze della veglia, codificate inizialmente nell'ippocampo, vengono durante il sonno riattivate attraverso le ripples, compresse e trasferite verso la corteccia attraverso il dialogo ippocampo-corticale orchestrato dagli spindles, e integrate nel patrimonio di conoscenze a lungo termine della corteccia attraverso le modificazioni sinaptiche indotte dalle onde lente.

Questo meccanismo è rilevante per il linguaggio in un modo che va al di là della semplice memorizzazione di parole nuove — fenomeno documentato e relativamente consolidato nella letteratura sul consolidamento linguistico nel sonno. La domanda più profonda e meno esplorata è: il sonno consolida anche la **struttura grammaticale** acquisita durante la veglia? E se sì, come si traduce questo consolidamento in termini di modifiche all'architettura oscillatoria corticale?

La risposta che si può avanzare, sulla base di un'integrazione degli studi disponibili, è la seguente. Il consolidamento grammaticale durante il sonno non riguarda la memorizzazione di regole esplicite — i parlanti nativi non hanno

accesso consapevole alle regole della propria grammatica — ma la **modificazione del paesaggio energetico degli attrattori oscillatori**: le strutture sintattiche incontrate durante la veglia, che hanno prodotto errori di previsione e aggiornamenti del modello generativo, vengono durante il sonno *riprocessate* attraverso il dialogo ippocampo-corticale, e le modificazioni sinaptiche che ne risultano approfondiscono i bacini di attrazione delle strutture grammaticali e amplificano le barriere tra attrattori corrispondenti a strutture diverse.

LXXVI. La formalizzazione del consolidamento notturno

La formalizzazione matematica di questo processo richiede di estendere il modello di apprendimento oscillatorio — già introdotto nella discussione sull'acquisizione — al dominio del sonno, dove le condizioni sono radicalmente diverse dalla veglia: nessun input linguistico esterno, oscillazioni di amplitudine molto maggiore, livelli neuromodulatori diversi (bassa acetilcolina, alta noradrenalina nelle prime fasi del sonno NREM).

Il meccanismo proposto è il seguente. Durante la veglia, l'elaborazione di una frase con struttura S produce un aggiornamento incrementale delle connessioni oscillatorie:

$$\Delta W_{\{ij\}}^{\{veglia\}}(S) \propto \lambda_{\{wake\}} \cdot \epsilon_i(S) \cdot \partial \theta_j / \partial W_{\{ij\}}$$

dove $\lambda_{\{wake\}}$ è la costante di apprendimento durante la veglia, relativamente piccola per evitare l'overfitting a singole istanze. L'insieme degli aggiornamenti accumulati durante una giornata di esposizione linguistica produce un vettore di modificazioni $\Delta W^{\{veglia\}}$ nello spazio dei pesi sinaptici.

Durante il sonno NREM, le onde lente corticali producono oscillazioni di ampiezza molto maggiore nello stesso spazio di stati delle reti oscillatorie corticali. Se i pesi sinaptici modificati durante la veglia hanno spostato leggermente il paesaggio energetico — abbozzando nuovi bacini di attrazione o approfondendo bacini preesistenti — le oscillazioni di grande ampiezza del sonno esplorano il paesaggio modificato con un'escursione molto più ampia. Questo processo, formalizzabile come una forma di **simulated annealing oscillatorio**, consente al sistema di convergere verso i minimi globali del paesaggio energetico anziché rimanere intrappolato nei minimi locali:

$$dW_{\{ij\}}/dt|_{\{sonno\}} = -\partial F_{\{eff\}}/\partial W_{\{ij\}} + \sqrt{2T_{\{eff\}}} \cdot \eta(t)$$

dove $F_{\{eff\}}$ è l'energia libera del sistema con i pesi modificati dalla veglia, $T_{\{eff\}}$ è una temperatura effettiva proporzionale all'ampiezza delle onde lente (elevata nelle prime ore del sonno profondo, decrescente con il progredire della

notte), e $\eta(t)$ è rumore stocastico. L'ampiezza elevata delle oscillazioni SWS corrisponde a un'alta temperatura effettiva che permette al sistema di attraversare barriere di potenziale altrimenti insuperabili, convergendo verso configurazioni di pesi che riflettono le regolarità grammaticali dell'input della veglia piuttosto che le peculiarità delle singole istanze.

L'effetto netto è che al mattino il paesaggio energetico grammaticale è più pronunciato di quanto fosse la sera: i bacini di attrazione delle strutture grammaticalmente frequenti sono più profondi, le barriere tra strutture diverse sono più alte, e la velocità di convergenza verso il corretto attrattore durante il parsing è aumentata. Questo corrisponde all'esperienza comune — e alla documentazione sperimentale — che le strutture apprese il giorno prima sono meglio padroneggiate il mattino seguente, e che la privazione del sonno compromette sistematicamente le capacità di elaborazione sintattica complessa.

LXXVII. Gli spindles talamici come selettori grammaticali

C'è un aspetto del consolidamento notturno che merita attenzione specifica: il ruolo degli **sleep spindles** nella selezione delle rappresentazioni da consolidare. Gli spindles non sono distribuiti uniformemente durante il sonno NREM: appaiono in raffiche, ciascuna della durata di uno–due secondi, e sono topograficamente organizzati — diversi tipi di spindles predominano in diverse regioni corticali, con gli spindles frontali (più lenti, ≈ 11 Hz) che riflettono attività prefrontale e gli spindles parietali (più rapidi, ≈ 13 – 15 Hz) che riflettono attività delle corteccie sensoriali e motorie.

La funzione selettiva degli spindles è documentata dal fatto che la loro densità durante il sonno è correlata con la quantità di apprendimento nella veglia precedente — in particolare con l'apprendimento di sequenze e di strutture regolari, più che con la memorizzazione di fatti arbitrari. Questa selettività è coerente con la proposta che gli spindles consolidino preferenzialmente le regolarità strutturali — le strutture grammaticali della lingua — piuttosto che le specificità lessicali delle frasi incontrate.

Il meccanismo proposto è il seguente: gli spindles talamici, generati dal nRT, attivano selettivamente i nuclei talamici specifici connessi alle aree corticali in cui le modificazioni sinaptiche della veglia sono state più pronunciate — le aree della rete del linguaggio dove l'elaborazione di strutture nuove o difficili ha prodotto i maggiori aggiornamenti. Questa attivazione selettiva amplifica le oscillazioni nelle reti corticali modificate dalla veglia, facilitando la riattivazione delle rappresentazioni sinapticamente rafforzate e la loro consolidazione attraverso le modificazioni sinaptiche indotte dagli spindles stessi (attraverso i canali del calcio voltaggio-dipendenti attivati dalla depolarizzazione spindle-indotta).

In termini del paesaggio energetico, gli spindles realizzano un'esplorazione *selettiva* del paesaggio — non casuale come quella delle onde lente, ma guidata dalla topografia delle modificazioni sinaptiche della veglia. Laddove le onde lente abbassano globalmente la temperatura effettiva, permettendo la convergenza verso i minimi globali, gli spindles dirigono l'esplorazione verso le regioni del paesaggio che la veglia ha reso più plastiche — combinando l'efficienza della convergenza globale con la specificità della selezione locale.

Questa architettura a due componenti — onde lente per la convergenza globale, spindles per la selezione locale — è la controparte notturna dell'architettura a due componenti della codifica predittiva diurna: le predizioni discendenti (beta) che organizzano globalmente la gerarchia, e i segnali di errore ascendenti (gamma) che correggono localmente le discrepanze. Il cervello, nella proposta oscillatoria, usa varianti della stessa architettura gerarchica sia durante la veglia che durante il sonno — con la differenza che di notte l'input esterno è assente e il sistema processa non il linguaggio del mondo ma il linguaggio accumulato nella propria memoria, riscrivendo il paesaggio energetico della grammatica con la lentezza e la profondità che la veglia non si può permettere.

Il sonno, in questa prospettiva, non è l'assenza di elaborazione linguistica: è la sua forma più silenziosa e più profonda — la notte in cui la grammatica si consolida, i bacini si approfondiscono, le barriere si alzano, e il sistema si prepara ad affrontare il linguaggio del giorno seguente con architetture oscillatorie un poco più robuste, un poco più precise, un poco più vicine alla competenza piena che il bambino costruisce nel corso di anni di veglie e di sonni alternati.

LXXVIII. Il silenzio sulla pragmatica e il suo peso

C'è un livello del linguaggio che l'intero impianto teorico di questo saggio ha affrontato sempre dall'esterno, senza mai penetrarne la struttura propria: il livello pragmatico. La pragmatica — il dominio che studia il significato in contesto, l'uso del linguaggio per compiere azioni, la relazione tra ciò che si dice e ciò che si intende — è rimasta ai margini della trattazione oscillatoria, affidata a formulazioni generiche sul "modello del discorso" e sulle "aspettative contestuali" senza ricevere la formalizzazione che gli altri livelli hanno ottenuto. Questo è un debito consistente, tanto più urgente quanto più si consideri che la pragmatica è il livello del linguaggio più direttamente connesso con la vita sociale e comunicativa dei parlanti — il livello in cui il linguaggio smette di essere un sistema formale e diventa uno strumento di azione nel mondo.

Il punto di partenza obbligato è la **teoria delle implicature conversazionali** di Herbert Paul Grice, formulata nelle *William James Lectures* del 1967 e pubblicata nel celebre saggio "Logic and Conversation" (1975). Grice aveva osservato che il significato trasmesso nel corso di una conversazione eccede sistematicamente il significato letterale delle frasi pronunciate — e che questo eccesso non è casuale

ma sistematico, regolato da principi che i partecipanti alla conversazione rispettano e su cui fanno affidamento. Il **Principio di Cooperazione** — *sii cooperativo nella conversazione* — si articola in quattro massime: quantità (*di' tanto quanto è necessario, non di più*), qualità (*di' solo ciò di cui sei ragionevolmente sicuro*), relazione (*sii pertinente*), e modo (*sii chiaro e ordinato*).

Le implicature nascono quando un parlante viola apparentemente una massima — producendo una risposta che, presa alla lettera, sembra poco informativa, irrilevante o eccessivamente vaga — e l'ascoltatore, presumendo che il parlante stia comunque rispettando il Principio di Cooperazione a un livello superiore, inferisce un significato aggiuntivo che rende la risposta cooperativa. Se a "Sei arrivato in orario?" si risponde "Ho trovato traffico", l'ascoltatore capisce che la risposta è "no" — non perché sia detto esplicitamente, ma perché la massima di relazione impone di attribuire alla risposta una pertinenza che richiede questa inferenza. L'implicatura è il significato indiretto calcolato attraverso questo ragionamento.

LXXIX. L'implicatura come aggiornamento del modello di discorso: formalizzazione bayesiana

La teoria di Grice è descrittiva e razionalista: descrive le implicature come il prodotto di un ragionamento esplicito sulle intenzioni del parlante. Ma l'esperienza quotidiana delle implicature suggerisce che esse non sono calcolate deliberatamente, passo per passo: vengono comprese automaticamente, rapidamente, spesso prima che l'inferenza esplicita sia completata. Questo suggerisce che l'elaborazione delle implicature sia parzialmente automatica — realizzata da processi che procedono in parallelo con l'elaborazione sintattica e semantica, senza richiedere deliberazione cosciente.

In termini della codifica predittiva, l'implicatura può essere formalizzata come un aggiornamento del modello generativo a un livello superiore rispetto a quello semantico: il livello del **modello dell'interlocutore** — la rappresentazione delle intenzioni, delle credenze e degli obiettivi del parlante che l'ascoltatore mantiene aggiornato nel corso della conversazione. Questo livello corrisponde, nella terminologia della Theory of Mind, alla rappresentazione mentalistica di primo o secondo ordine dell'interlocutore.

Formalmente, si può estendere il modello generativo gerarchico introdotto nella parte IV aggiungendo un livello superiore che rappresenta lo stato mentale dell'interlocutore:

$$p(o, s, m, c, I) = p(o \mid s) \cdot p(s \mid m, c) \cdot p(m \mid c, I) \cdot p(c \mid I) \cdot p(I)$$

dove I è lo stato intenzionale dell'interlocutore — le sue credenze, i suoi obiettivi comunicativi, le sue aspettative sulla risposta dell'ascoltatore. Il calcolo dell'implicatura corrisponde all'inferenza su I data l'osservazione di o : perché mai un parlante cooperativo produrrebbe *questa* risposta in *questo* contesto, se non per comunicare qualcosa di specifico su I ? La risposta ottimale — il valore di I che massimizza la probabilità *a posteriori* data l'osservazione — determina l'implicatura calcolata.

Questo schema bayesiano è formalmente equivalente al ragionamento griceano, ma è implementato come minimizzazione dell'energia libera variazionale nel modello generativo — un processo che può avvenire in parallelo con l'elaborazione sintattico-semantica, senza richiedere deliberazione sequenziale. Le implicature vengono "calcolate" nel senso in cui il cervello "calcola" qualunque inferenza predittiva: attraverso la convergenza iterativa della gerarchia di oscillatori verso lo stato che minimizza l'energia libera a tutti i livelli, incluso il livello della rappresentazione dell'interlocutore.

LXXX. Gli atti linguistici e la dimensione performativa del linguaggio

La teoria degli **atti linguistici** — sviluppata da John Langshaw Austin nelle *How to Do Things with Words* (le William James Lectures del 1955, pubblicate postume nel 1962) e sistematizzata da John Searle in *Speech Acts* (1969) — aveva portato all'attenzione della filosofia del linguaggio una dimensione che la semantica formale tendeva a ignorare: il fatto che il linguaggio non sia soltanto un mezzo per descrivere stati di cose nel mondo, ma uno strumento per *fare* cose — promettere, ordinare, dichiarare, affermare, scusarsi, sposarsi. Austin aveva distinto tra **atti locutori** (il dire qualcosa con un certo significato), **atti illocutori** (il fare qualcosa nel dire — promettere, ordinare, affermare) e **atti perlocutori** (il produrre un certo effetto nell'ascoltatore attraverso il dire).

La dimensione performativa del linguaggio è quella che più direttamente sfida le teorie computazionali del processamento linguistico: come fa il cervello a determinare che "Puoi passarmi il sale?" è una richiesta e non una domanda sulla capacità fisica dell'ascoltatore? Come determina che "Il battesimo di questo vascello è dichiarato aperto" è un atto che cambia lo stato del mondo — non descrive, crea? Queste determinazioni richiedono non soltanto l'elaborazione del contenuto proposizionale della frase, ma la sua integrazione con il contesto istituzionale, con le convenzioni sociali, con il modello delle intenzioni dell'interlocutore — esattamente il livello superiore della gerarchia predittiva che si è introdotto nella sezione precedente.

In termini oscillatori, la comprensione della forza illocutoria di un atto linguistico richiede che le oscillazioni infra-lente della rete di default — il livello del modello contestuale e intersoggettivo — modulino in modo specifico le oscillazioni della gerarchia sintattico-semantica. Il riconoscimento che "Puoi passarmi il sale?" è

una richiesta non emerge dall'analisi sintattica della frase (che è grammaticalmente una domanda) né dall'analisi semantica del suo contenuto proposizionale (che concerne la capacità dell'ascoltatore), ma dalla modulazione di queste analisi da parte del modello contestuale: in questo contesto, con questo interlocutore, con questi obiettivi condivisi, la frase ha forza illocutoria di richiesta. Questa modulazione è, nella proposta oscillatoria, una forma di coupling tra le oscillazioni infra-lente del modello contextuale (< 0.5 Hz) e le oscillazioni delta-theta della struttura sintattico-semantica — il livello più alto della gerarchia oscillatoria che dialoga con il livello sintattico attraverso connessioni discendenti che non sono state ancora sistematicamente studiate.

LXXXI. L'ironia e la doppia codifica oscillatoria

L'ironia — il caso in cui si intende il contrario di ciò che si dice — è forse l'esempio più istruttivo della dissociazione tra significato letterale e significato comunicativo, e al tempo stesso uno dei fenomeni pragmatici più difficili da gestire per le teorie computazionali del linguaggio. Comprendere l'ironia richiede di elaborare il significato letterale della frase, rilevare la sua incongruenza con il contesto, inferire che l'incongruenza è intenzionale, e calcolare il significato inteso come inversione o negazione del significato letterale — tutto questo in un tempo dell'ordine di centinaia di millisecondi, spesso prima che la frase sia completata.

Le evidenze neuroimaging sull'ironia mostrano sistematicamente un maggiore coinvolgimento dell'emisfero destro rispetto alla comprensione letterale — in particolare del giro frontale inferiore destro e della corteccia temporale superiore posteriore destra. Questo risultato è coerente con la proposta di Poeppel sull'asimmetria emisferica: l'emisfero destro, specializzato per l'elaborazione a grana grossa e per la finestra di integrazione delta, è quello che mantiene il contesto discorsivo ampio necessario per rilevare l'incongruenza tra il significato letterale della frase e il contesto in cui essa è prodotta.

In termini oscillatori, la comprensione dell'ironia richiederebbe una **doppia codifica**: un primo percorso di elaborazione che processa il significato letterale attraverso la gerarchia sintattico-semantica sinistra (theta-beta nelle corteccie temporali e prefrontali sinistre), e un secondo percorso che elabora l'incongruenza contestuale attraverso le oscillazioni delta dell'emisfero destro e le connessioni con il modello discorsivo-intersoggettivo. L'integrazione delle due codifiche — il riconoscimento che il significato letterale è intenzionalmente incongruente con il contesto, e dunque il calcolo del significato ironico come l'inversione del letterale — richiederebbe un accoppiamento tra le oscillazioni delta destre (contesto) e le oscillazioni theta-beta sinistre (significato letterale), mediato probabilmente dalle connessioni callosali tra i due emisferi.

Questa proposta è compatibile con la maggiore latenza della comprensione ironica rispetto a quella letterale, documentata sperimentalmente: la doppia codifica e la sua integrazione attraverso il corpo calloso richiedono un tempo di elaborazione aggiuntivo rispetto alla semplice comprensione letterale, che può avvenire attraverso la gerarchia sinistra senza il coinvolgimento della doppia codifica emisferica.

LXXXII. La conversazione come accoppiamento tra sistemi oscillatori distinti

C'è una dimensione della pragmatica che nessuna teoria del processamento linguistico individuale può catturare: il fatto che la conversazione non sia elaborazione di un segnale da parte di un sistema, ma un processo di **coordinazione dinamica tra due o più sistemi oscillatori distinti** — i cervelli dei partecipanti alla conversazione. La proposta oscillatoria non è soltanto una teoria del processamento linguistico individuale: è potenzialmente una teoria della coordinazione intersoggettiva, del modo in cui due cervelli si sincronizzano per produrre la comprensione condivisa che la conversazione richiede.

Studi di **iper-scanning** — la misurazione simultanea dell'attività cerebrale di due o più partecipanti durante una conversazione naturale, attraverso EEG o fMRI simultanei — hanno documentato fenomeni di sincronizzazione intercerebrale (*brain-to-brain synchrony*) durante la comunicazione. Le oscillazioni theta delle corteccie frontali del parlante e dell'ascoltatore si sincronizzano durante la conversazione naturale, con un ritardo compatibile con il ritardo di trasmissione del segnale acustico tra i due partecipanti. Questa sincronizzazione è assente durante l'ascolto di parlato registrato — quando cioè la comunicazione è unidirezionale e non interattiva — e si riduce quando la comprensione si deteriora.

La sincronizzazione intercerebrale è, nella proposta oscillatoria, la firma neurofisiologica del **comune terreno** (*common ground*) che i partecipanti alla conversazione costruiscono e mantengono: la convergenza verso stati oscillatori condivisi è il correlato dinamico della costruzione di un modello del discorso condiviso, di una rappresentazione comune degli eventi, dei referenti e delle intenzioni che permette la coordinazione comunicativa. La conversazione non è uno scambio di messaggi tra sistemi separati: è la costruzione progressiva di un sistema accoppiato — un sistema dinamico distribuito tra i cervelli dei partecipanti — che converge verso stati di sincronizzazione condivisa.

Questo è il livello più alto della gerarchia oscillatoria del linguaggio — il livello che nessuno studio del singolo cervello in isolamento può misurare — e al tempo stesso il livello in cui la proposta oscillatoria converge con le tradizioni filosofiche che hanno insistito sulla natura fondamentalmente intersoggettiva del

linguaggio: la concezione wittgensteiniana del linguaggio come pratica condivisa, la pragmatica della comunicazione di Austin e Searle, la teoria dell'azione comunicativa di Habermas. La sincronizzazione oscillatoria intercerebrale non è una metafora di queste intuizioni: è il meccanismo biofisico che le realizza, la forma in cui la natura ha risolto il problema di come due sistemi biologici distinti convergano verso la comprensione condivisa che il linguaggio rende possibile — e che il linguaggio, al tempo stesso, presuppone.

Il linguaggio è, in ultima analisi, questo: due cervelli che oscillano, che si accoppiano, che costruiscono insieme uno spazio di significato condiviso che non esisteva prima che iniziassero a parlare. La struttura gerarchica della sintassi, realizzata nell'architettura temporale delle oscillazioni, è il mezzo con cui questo accoppiamento avviene in modo preciso e ricco — non un generico contagio di stati emotivi, ma una trasmissione strutturata di contenuto proposizionale, intenzione illocutoria e orientamento pragmatico. Ed è in questo accoppiamento — nella sincronia tra sistemi oscillatori che rimangono distinti pur coordinandosi — che il linguaggio mostra la propria natura più profonda: non una proprietà del singolo cervello, ma una proprietà dello spazio tra i cervelli, realizzata *nei* cervelli attraverso le oscillazioni che li attraversano.

Appendice Tecnica A: Derivazioni matematiche

A.1 La trasformata di Hilbert e il segnale analitico

La decomposizione ampiezza-fase di un segnale reale $x(t)$ — il punto di partenza di qualunque misura di accoppiamento cross-frequenza — si basa sulla costruzione del **segnale analitico** $z(t)$, definito come la parte non-ridondante della rappresentazione complessa del segnale. La definizione precisa della trasformata di Hilbert come integrale di valor principale:

$$\hat{x}(t) = H\{x\}(t) = (1/\pi) \cdot \text{P.V.} \int_{-\infty}^{+\infty} x(\tau)/(t - \tau) d\tau$$

può essere derivata a partire dal requisito che lo spettro di $z(t) = x(t) + i \cdot \hat{x}(t)$ sia nullo per frequenze negative. Nello spazio delle frequenze, l'operatore di Hilbert corrisponde alla moltiplicazione per $-i \cdot \text{sgn}(f)$:

$$\mathcal{F}\{H\{x\}\}(f) = -i \cdot \text{sgn}(f) \cdot X(f)$$

dove $X(f) = \mathcal{F}\{x\}(f)$ è la trasformata di Fourier di $x(t)$ e $\text{sgn}(f) = +1$ per $f > 0$, -1 per $f < 0$. Il segnale analitico ha dunque spettro:

$$Z(f) = X(f) + i \cdot (-i \cdot \text{sgn}(f) \cdot X(f)) = X(f) \cdot (1 + \text{sgn}(f)) = 2 \cdot X(f) \cdot u(f)$$

dove $u(f)$ è la funzione gradino di Heaviside. Lo spettro di $z(t)$ è dunque il doppio dello spettro di $x(t)$ per $f > 0$ e zero per $f < 0$ — esattamente il requisito di analiticità.

La fase istantanea $\varphi(t)$ e l'ampiezza istantanea $A(t)$ si ottengono come:

$$A(t) = |z(t)| = \sqrt{x(t)^2 + \hat{x}(t)^2}$$

$$\varphi(t) = \arg(z(t)) = \arctan(\hat{x}(t) / x(t))$$

Affinché queste quantità abbiano significato fisico come "fase" e "ampiezza" di un'oscillazione quasi-monocromatica, è necessario che il segnale sia a banda stretta in prossimità della frequenza di interesse. Questa condizione è garantita dal filtraggio passa-banda applicato prima della trasformata di Hilbert: dato un segnale EEG grezzo $x(t)$, si estrae la componente nella banda $[f_L, f_H]$ attraverso un filtro lineare con risposta in frequenza $B(f)$:

$$x_{\text{band}}(t) = \mathcal{F}^{-1}\{B(f) \cdot X(f)\}$$

e si applica la trasformata di Hilbert a $x_{\text{band}}(t)$. La scelta del filtro — ordine, risposta di fase, larghezza di banda — influenza significativamente le misure di PAC risultanti, ed è una delle principali fonti di variabilità metodologica tra studi diversi. I filtri a fase zero (zero-phase filters, come il filtro di Butterworth applicato in avanti e indietro con `filtfilt`) sono preferibili per evitare distorsioni di fase che altererebbero le misure di accoppiamento.

A.2 Derivazione completa del Modulation Index

Il Modulation Index (MI) di Tort e collaboratori è una misura della divergenza tra la distribuzione dell'ampiezza dell'oscillazione rapida rispetto alla fase dell'oscillazione lenta e la distribuzione uniforme attesa in assenza di accoppiamento.

Dati il segnale di fase lenta $\varphi_L(t)$ e il segnale di ampiezza rapida $A_H(t)$, si divide il ciclo $[-\pi, \pi]$ in N bin di uguale ampiezza $\Delta\varphi = 2\pi/N$. Per ogni bin $k \in \{1, \dots, N\}$, si calcola l'ampiezza media dell'oscillazione rapida negli istanti in cui $\varphi_L(t)$ cade nel k -esimo bin:

$$\langle A_H \rangle_k = (1/n_k) \cdot \sum_{t: \varphi_L(t) \in \text{bin}_k} A_H(t)$$

dove n_k è il numero di campioni nel k -esimo bin. Si normalizza per ottenere una distribuzione di probabilità discreta:

$$p_k = \langle A_H \rangle_k / \sum_{j=1}^N \langle A_H \rangle_j, \quad \text{con } \sum_{k=1}^N p_k = 1$$

La divergenza di Kullback-Leibler dalla distribuzione uniforme $U_k = 1/N$ è:

$$\begin{aligned} D_{\{KL\}}(p \parallel U) &= \sum_{k=1}^N p_k \cdot \log(p_k / U_k) = \sum_{k=1}^N p_k \cdot \log(N \cdot p_k) \\ &= \log(N) + \sum_{k=1}^N p_k \cdot \log(p_k) = \log(N) - H(p) \end{aligned}$$

dove $H(p) = -\sum_k p_k \cdot \log(p_k)$ è l'entropia di Shannon della distribuzione p . Si verifica che $H(p) \in [0, \log(N)]$, con $H(p) = \log(N)$ se e solo se p è uniforme, e $H(p) = 0$ se e solo se p è concentrata su un singolo bin. Il Modulation Index è:

$$MI = D_{\{KL\}}(p \parallel U) / \log(N) = 1 - H(p)/\log(N)$$

Alcune proprietà rilevanti:

$$\begin{aligned} MI = 0 &\Leftrightarrow p = U \quad (\text{distribuzione uniforme, nessun accoppiamento}) \\ MI = 1 &\Leftrightarrow p_k = \delta_{k,k^*} \quad \text{per qualche } k^* \quad (\text{accoppiamento massimo}) \\ MI &\in [0, 1] \quad \text{per ogni distribuzione } p \quad (\text{dalla disuguaglianza } H(p) \leq \log(N)) \end{aligned}$$

La scelta $N = 18$ (bin di 20°) è un compromesso empirico tra risoluzione della distribuzione di fase e stabilità statistica della stima: con N troppo grande e finestre temporali brevi, le stime $\langle A_H \rangle_k$ diventano instabili per insufficienza di campioni per bin. La significatività statistica del MI viene stabilita attraverso permutazioni: si genera una distribuzione nulla permutando i trial o traslando la fase del segnale lento di un intervallo casuale rispetto al segnale rapido, e si calcola il percentile del MI osservato nella distribuzione nulla.

A.3 Analisi di stabilità del modello di Kuramoto e derivazione della transizione critica

Il modello di Kuramoto per N oscillatori con accoppiamento globale è:

$$d\theta_i/dt = \omega_i + (K/N) \cdot \sum_{j=1}^N \sin(\theta_j - \theta_i), \quad i = 1, \dots, N$$

Introducendo il parametro d'ordine complesso $r \cdot e^{i\psi} = (1/N) \sum_j e^{i\theta_j}$, le equazioni si riscrivono come:

$$d\theta_i/dt = \omega_i + K \cdot r \cdot \sin(\psi - \theta_i)$$

Nel limite termodinamico $N \rightarrow \infty$, con frequenze naturali distribuite secondo $g(\omega)$ (densità di probabilità simmetrica attorno a ω_0 con larghezza caratteristica γ), si cerca la soluzione auto-consistente stazionaria in un sistema di riferimento rotante a frequenza $\psi = \omega_0$:

$$d\varphi_i/dt = \nu_i + K \cdot r \cdot \sin(-\varphi_i) \quad \text{con} \quad \nu_i = \omega_i - \omega_0$$

Gli oscillatori si dividono in due popolazioni: gli **oscillatori bloccati** ($|\nu_i| < K \cdot r$), che si sincronizzano a una fase stazionaria φ_i^* tale che $\sin(\varphi_i^*) = \nu_i / (K \cdot r)$; e gli **oscillatori liberi** ($|\nu_i| > K \cdot r$), che ruotano a frequenze diverse e il cui contributo al parametro d'ordine si media a zero.

La condizione di auto-consistenza per r si ottiene calcolando il contributo degli oscillatori bloccati:

$$r = \int_{-\infty}^{\infty} \{ |\nu| < Kr \} \cos(\varphi^*(\nu)) \cdot g(\nu + \omega_0) d\nu$$

dove $\cos(\varphi^*(\nu)) = \sqrt{1 - (\nu/Kr)^2}$ è la componente in fase dell'oscillatore bloccato con desintonizzazione ν . Cambiando variabile $\nu = K \cdot r \cdot \sin(\varphi)$:

$$r = K \cdot r \cdot \int_{-\pi/2}^{\pi/2} \cos^2(\varphi) \cdot g(K \cdot r \cdot \sin(\varphi) + \omega_0) d\varphi$$

Per la distribuzione lorentziana $g(\omega) = (\gamma/\pi) / ((\omega - \omega_0)^2 + \gamma^2)$, questa integrale si calcola per residui, dando:

$$1 = K \cdot \int_{-\pi/2}^{\pi/2} \cos^2(\varphi) \cdot g(K \cdot r \cdot \sin(\varphi)) d\varphi = K/(2\gamma) \quad \text{per } r \rightarrow 0^+$$

L'equazione $1 = K/(2\gamma)$ determina il valore critico dell'accoppiamento:

$$K_c = 2\gamma$$

Per $K > K_c$, si cerca la soluzione $r > 0$ dell'equazione di auto-consistenza. Per K vicino a K_c ($K = K_c + \epsilon$, con ϵ piccolo), si espande $g(K \cdot r \cdot \sin(\varphi))$ attorno a $K \cdot r = 0$:

$$g(K \cdot r \cdot \sin(\varphi)) \approx g(0) - (K \cdot r \cdot \sin(\varphi))^2 \cdot g''(0)/2 + \dots$$

Sostituendo nell'equazione di auto-consistenza e raccogliendo i termini in r :

$$r^2 \approx (4\gamma/\pi) \cdot (K - K_c) / K_c^2$$

per la distribuzione lorentziana, che dà:

$$r \propto \sqrt{K - K_c} \quad \text{per } K \rightarrow K_c^+$$

confermando la natura di transizione di fase del secondo ordine con esponente critico $\beta = 1/2$.

L'**energia libera** del sistema di Kuramoto (il potenziale di Lyapunov nella forma standard):

$$V[\{\theta_i\}] = -(K/2N) \cdot \sum_{i,j} \cos(\theta_j - \theta_i)$$

è minimizzata nella configurazione sincrona $\theta_i = \psi$ per tutti gli i , e il minimo è:

$$V_{\min} = -K \cdot N \cdot r^2/2$$

La differenza di energia libera tra lo stato incoerente ($r = 0$) e quello coerente ($r > 0$) è:

$$\Delta V = V(r=0) - V(r>0) = K \cdot N \cdot r^2/2 > 0 \quad \text{per } K > K_c$$

che conferma che lo stato coerente è il minimo energetico globale al di sopra della transizione.

A.4 Il modello gerarchico di Kuramoto con accoppiamento top-down: condizioni di sincronizzazione n:m

Nel modello gerarchico, un oscillatore "lento" con fase $\Theta(t)$ e frequenza Ω modula gli oscillatori "rapidi" con fasi $\theta_i(t)$ e frequenze ω_i attraverso un termine di accoppiamento top-down:

$$d\theta_i/dt = \omega_i + K_{\text{fast}} \cdot r_{\text{fast}} \cdot \sin(\psi_{\text{fast}} - \theta_i) + K_{\text{top}} \cdot \sin(n \cdot \Theta - m \cdot \theta_i)$$

Il termine $K_{\text{top}} \cdot \sin(n \cdot \Theta - m \cdot \theta_i)$ produce un accoppiamento n:m (Arnold tongue) tra l'oscillatore lento e ciascun oscillatore rapido. Per K_{top} sufficientemente grande, gli oscillatori rapidi vengono trascinati verso la relazione di fase:

$$m \cdot \theta_i \approx n \cdot \Theta + \text{const}_i$$

ossia m cicli dell'oscillatore rapido corrispondono a n cicli dell'oscillatore lento. La condizione di risonanza $n:m$ è soddisfatta quando:

$$|m \cdot \omega_i - n \cdot \Omega| < K_{\text{top}}$$

che definisce la **lingua di Arnold** nello spazio dei parametri $(\Omega/\omega_i, K_{\text{top}}/\omega_i)$: le coppie di frequenze per cui la sincronizzazione $n:m$ è stabile. La lingua più ampia è quella 1:1 (sincronizzazione a fase identica); le lingue 1:2, 1:3, 2:3 etc. hanno ampiezze decrescenti come K_{top}^{n+m} nell'approssimazione di accoppiamento debole.

Per la gerarchia oscillatoria del linguaggio, i rapporti $n:m$ rilevanti sono determinati dai rapporti empirici tra le frequenze caratteristiche delle bande: il rapporto $\theta/\delta \approx 3-5:1$ suggerisce accoppiamenti 3:1, 4:1 o 5:1 tra δ e θ ; il rapporto $\gamma/\theta \approx 5-10:1$ suggerisce accoppiamenti 5:1 fino a 10:1 tra θ e γ . La presenza di accoppiamenti $n:m$ multipli produce una struttura di sincronizzazione quasi-periodica che può essere analizzata attraverso la teoria delle perturbazioni della curva di risposta di fase (PRC — Phase Response Curve):

$$\Delta\varphi_i = Z(\varphi_i) \cdot I(t)$$

dove $Z(\varphi)$ è la PRC dell'oscillatore (la variazione di fase in risposta a un impulso unitario in funzione della fase corrente) e $I(t)$ è il segnale di accoppiamento. Per oscillatori del tipo quasi-sinusoidale (vicini alla biforcazione di Hopf), la PRC ha forma sinusoidale $Z(\varphi) \approx \sin(\varphi)$, e l'analisi del metodo della media (averaging method) sulla dinamica della differenza di fase $\psi = n \cdot \Theta - m \cdot \theta_i$ dà:

$$\begin{aligned} d\psi/dt &= n \cdot \Omega - m \cdot \omega_i - m \cdot K_{\text{top}} \cdot \langle Z(\varphi_i) \cdot \sin(n \cdot \Theta - m \cdot \varphi_i) \rangle \\ &\approx \Delta\omega - (m \cdot K_{\text{top}}/2) \cdot \sin(\psi) \end{aligned}$$

dove $\Delta\omega = n \cdot \Omega - m \cdot \omega_i$ è la desintonizzazione e il termine di media produce l'equazione di Adler standard, la cui soluzione stazionaria $\psi^* = \arcsin(2\Delta\omega / (m \cdot K_{\text{top}}))$ esiste se e solo se $|\Delta\omega| < m \cdot K_{\text{top}}/2$.

A.5 Energia libera variazionale: derivazione della decomposizione bound-surprise

Partendo dalla definizione della divergenza KL e dalla sua relazione con la verosimiglianza marginale, si mostra la derivazione completa della decomposizione dell'energia libera.

Per qualunque distribuzione $q(s)$:

$$\begin{aligned}
 \log p(o) &= \log \int p(o, s) \, ds \\
 &= \log \int q(s) \cdot [p(o, s)/q(s)] \, ds \\
 &\geq \int q(s) \cdot \log[p(o, s)/q(s)] \, ds \quad (\text{disuguaglianza di Jensen, poiché } \log \text{ è concavo}) \\
 &= \mathbb{E}_q[\log p(o, s)] - \mathbb{E}_q[\log q(s)] \\
 &= \mathbb{E}_q[\log p(o | s)] + \mathbb{E}_q[\log p(s)] - \mathbb{E}_q[\log q(s)] \\
 &= \mathbb{E}_q[\log p(o | s)] - D_{\text{KL}}(q(s) \parallel p(s))
 \end{aligned}$$

Definendo l'**energia libera variazionale** come:

$$F[q] = -\mathbb{E}_q[\log p(o | s)] + D_{\text{KL}}(q(s) \parallel p(s))$$

si ha:

$$\log p(o) = -F[q] + D_{\text{KL}}(q(s) \parallel p(s | o)) \geq -F[q]$$

poiché $D_{\text{KL}} \geq 0$. Dunque $F[q] \geq -\log p(o) = \text{sorpresa}(o)$, con uguaglianza sse $q(s) = p(s | o)$. Minimizzare F rispetto a q equivale a:

- (a) avvicinare q alla vera posteriore $p(s | o)$ [termine KL],
- (b) massimizzare la verosimiglianza attesa $\mathbb{E}_q[\log p(o | s)]$ [termine di accuracy],
- (c) minimizzare la complessità $D_{\text{KL}}(q \parallel p)$ [termine di complexity, penalità sull'allontanamento dal prior].

La decomposizione alternativa in termini di **accuracy** e **complexity** è:

$$\begin{aligned}
 F[q] &= -\mathbb{E}_q[\log p(o | s)] + D_{\text{KL}}(q(s) \parallel p(s)) \\
 &= -\text{accuracy} + \text{complexity}
 \end{aligned}$$

che mostra esplicitamente il trade-off fondamentale dell'inferenza bayesiana approssimata: una distribuzione q troppo distante dal prior ha alta complexity; una distribuzione q troppo fedele al prior ha bassa accuracy. Il minimo di F bilancia questi due termini.

Per la forma gaussiana $q(s) = \mathcal{N}(\mu, \Sigma)$ e likelihood lineare-gaussiana $\log p(o | s) = -(1/2\sigma^2) | |o - Cs| |^2$, la minimizzazione di F rispetto a μ e Σ dà le equazioni di aggiornamento del filtro di Kalman — confermando che la codifica predittiva gaussiana è un caso speciale di inferenza variazionale con likelihood e prior gaussiani.

A.6 Transfer Entropy come misura di causalità oscillatoria

La Transfer Entropy (TE) — introdotta da Schreiber nel 2000 come misura non-parametrica della causalità direzionale tra processi stocastici — è particolarmente rilevante per distinguere il flusso di informazione ascendente (errori di previsione, gamma) da quello discendente (predizioni, beta) nelle reti oscillatorie corticali.

Data una coppia di processi stocastici $X(t)$ e $Y(t)$, la Transfer Entropy da Y a X è:

$$T_{Y \rightarrow X} = \sum_{x_{t+1}, x_t^{(k)}, y_t^{(l)}} p(x_{t+1}, x_t^{(k)}, y_t^{(l)}) \cdot \log \left[\frac{p(x_{t+1} | x_t^{(k)}, y_t^{(l)})}{p(x_{t+1} | x_t^{(k)})} \right]$$

dove $x_t^{(k)} = (x_t, x_{t-1}, \dots, x_{t-k+1})$ e $y_t^{(l)} = (y_t, y_{t-1}, \dots, y_{t-l+1})$ sono i vettori di storia di lunghezza k e l rispettivamente. La TE è la riduzione dell'incertezza su x_{t+1} ottenuta conoscendo la storia di Y oltre a quella di X : misura dunque il contributo causale direzionale di Y su X .

La **TE netta** (net Transfer Entropy) o **TE asimmetrica**:

$$\Delta T_{Y \rightarrow X} = T_{Y \rightarrow X} - T_{X \rightarrow Y}$$

misura la predominanza del flusso causale in una direzione rispetto all'altra. Un valore $\Delta T_{Y \rightarrow X} > 0$ indica che Y causa X più di quanto X causi Y — corrispondente, nella proposta oscillatoria, a una connessione discendente netta (se Y è la regione superiore e X quella inferiore) o a una connessione ascendente netta (nel caso inverso).

L'applicazione della TE ai segnali EEG banda-filtrati richiede la stima della distribuzione di probabilità congiunta — tipicamente attraverso k -nearest neighbors o binning adattivo — e la correzione per i bias di stima finita attraverso permutazioni o tecniche bootstrap. La risoluzione spaziale è limitata dalla natura volumetrica dell'EEG; studi ECoG e MEG permettono stime più affidabili della direzionalità del flusso nelle reti corticali.

Appendice Tecnica B: Metodologia sperimentale e paradigmi

B.1 Le tecniche di misura delle oscillazioni corticali nel linguaggio

La scelta della tecnica di misura condiziona in modo fondamentale il tipo di domande empiriche che possono essere poste alla proposta oscillatoria. Le principali tecniche disponibili — EEG, MEG, ECoG, fMRI — hanno profili di risoluzione spazio-temporale complementari che le rendono adatte a diversi aspetti della teoria.

L'**elettroencefalografia** (EEG) misura le fluttuazioni del potenziale elettrico sulla superficie cranica, originate dai potenziali postsinaptici delle popolazioni di neuroni piramidali corticali orientati perpendicolarmente alla superficie. La risoluzione temporale è dell'ordine del millisecondo — superiore a qualunque altra tecnica non invasiva — rendendola ideale per lo studio delle oscillazioni e del loro accoppiamento cross-frequenza. La risoluzione spaziale è tuttavia scarsa (dell'ordine dei centimetri) a causa della dispersione del segnale attraverso il fluido cerebrospinale, le meningi e il cranio, e dell'ambiguità del problema inverso (ricavare le sorgenti dalla distribuzione di potenziale superficiale). I sistemi EEG ad alta densità (128–256 elettrodi) e le tecniche di analisi della sorgente (LORETA, beamforming) migliorano la localizzazione spaziale, ma rimangono intrinsecamente meno precisi della MEG.

La **magnetoencefalografia** (MEG) misura i campi magnetici generati dalle stesse correnti neuronali che producono il segnale EEG. I campi magnetici hanno il vantaggio di non essere distorti dal cranio e dalle meningi — a differenza dei campi elettrici — il che migliora significativamente la risoluzione spaziale (dell'ordine del centimetro o meno per sorgenti superficiali) pur mantenendo la stessa risoluzione temporale del millisecondo. La MEG è particolarmente sensibile alle sorgenti nelle solcature corticali, orientate tangenzialmente alla superficie — mentre l'EEG è più sensibile alle sorgenti nella girasazione, orientate radialmente. La complementarità delle due tecniche rende la combinazione EEG-MEG particolarmente potente per la localizzazione delle sorgenti oscillatorie.

L'**elettrocorticografia** (ECoG) è una tecnica invasiva che richiede l'impianto di elettrodi direttamente sulla superficie corticale o nella profondità del tessuto cerebrale, attraverso procedure neurochirurgiche. La risoluzione spazio-temporale è nettamente superiore a EEG e MEG: la risoluzione spaziale è dell'ordine del millimetro (distanza tra elettrodi adiacenti nelle griglie epidurali) e la risoluzione temporale rimane dell'ordine del millisecondo. L'ECoG è praticabile soltanto su pazienti che vengono sottoposti a neurochirurgia per ragioni cliniche indipendenti (epilessia farmacoresistente, tumori corticali), e la copertura anatomica è limitata alle regioni implantate per ragioni cliniche. Nonostante questi limiti, l'ECoG ha fornito alcune delle evidenze più dirette e convincenti sull'accoppiamento theta-gamma durante l'elaborazione linguistica, e rimane la tecnica di elezione per rispondere a domande sulla dinamica temporale fine delle reti corticali.

L'**imaging per risonanza magnetica funzionale** (fMRI) misura le variazioni del segnale BOLD (*Blood-Oxygen-Level-Dependent*), un correlato indiretto

dell'attività neurale attraverso il coupling neuro-vascolare. La risoluzione spaziale è ottima (dell'ordine del millimetro per sistemi ad alto campo), ma la risoluzione temporale è dell'ordine dei secondi — molto più lenta del millisecondo delle oscillazioni neurali. L'fMRI non è dunque adatta alla misura diretta delle oscillazioni e del loro accoppiamento, ma è preziosa per la localizzazione anatomica delle reti attivate durante l'elaborazione linguistica e per la mappatura della connettività funzionale a riposo — la struttura di correlazioni temporali tra regioni corticali distanti che riflette la connettività anatomica sottostante.

B.2 Paradigmi sperimentali per la misurazione della gerarchia oscillatoria sintattica

La misurazione della firma oscillatoria della struttura sintattica richiede paradigmi che separino i contributi della struttura dalla struttura prosodica, ritmica, lessicale e semantica — operazione non banale, poiché questi aspetti co-variano nel linguaggio naturale. I paradigmi più influenti sono i seguenti.

Il paradigma delle frasi con dipendenze a lungo raggio vs. breve raggio presenta ai partecipanti frasi che differiscono sistematicamente nella distanza strutturale tra elementi dipendenti — tipicamente soggetto e verbo principale, o verbo e il suo argomento — mantenendo costanti lunghezza della frase, frequenza lessicale, e struttura prosodica. Esempio: *"Il bambino ha mangiato la torta"* (dipendenza breve) vs. *"Il bambino che la maestra aveva rimproverato ha mangiato la torta"* (dipendenza lunga attraverso una relativa). La differenza nei pattern di PAC tra le due condizioni — misurata nelle bande beta e theta delle cortecce frontali e temporali sinistre — è la firma oscillatoria del mantenimento della dipendenza strutturale.

Il paradigma delle frasi anomale confronta frasi grammaticalmente corrette con frasi contenenti violazioni sintattiche selettive — tipicamente anomalie di accordo morfologico, di ordine sintagmatico, o di selezione di categoria lessicale — mantenendo costante la struttura prosodica e il contenuto lessicale. Le frasi anomale producono ERP caratteristici (ELAN, P600) e si predice che producano pattern di PAC alterati — in particolare riduzioni del MI nelle bande theta-beta che realizzano la dipendenza sintattica violata.

Il paradigma RSVP (*Rapid Serial Visual Presentation*) presenta le parole della frase una alla volta su schermo, a velocità controllata, eliminando la struttura prosodica dell'input acustico. Questo paradigma permette di dissociare la struttura sintattica dall'intonazione, ma elimina anche il ritmo naturale del parlato che guida l'entrainment oscillatorio. I pattern oscillatori durante la RSVP sono necessariamente endogeni — generati dalla gerarchia corticale senza un segnale esterno con cui sincronizzarsi — e dunque più direttamente attribuibili alla struttura sintattica che al ritmo acustico.

Il **paradigma delle frasi Jabberwocky** — ispirato al celebre poema di Lewis Carroll — usa frasi con struttura sintattica intatta ma con parole di contenuto sostituite da non-parole fonologicamente plausibili: *"Il blicco ha granfato la torba"*. Le parole funzionali (articoli, preposizioni, morfemi flessivi) sono mantenute reali, preservando la struttura sintattica; le parole di contenuto (nomi, verbi, aggettivi) sono sostituite con non-parole, eliminando il contenuto semantico lessicale. Confrontare le risposte oscillatorie a frasi Jabberwocky con quelle a frasi normali permette di isolare la componente oscillatoria specificamente associata alla struttura sintattica da quella associata all'elaborazione semantica lessicale.

Il **paradigma di iper-scanning** per la conversazione diadica misura le oscillazioni di due partecipanti simultaneamente mentre conversano naturalmente, alternando i ruoli di parlante e ascoltatore. Le misure di sincronizzazione intercerebrale — PLV o coerenza calcolati tra segnali EEG/MEG dei due partecipanti — e la loro variazione in funzione del contenuto e della struttura della conversazione (turni, dipendenze discorsive, atti illocutori) forniscono una finestra sulla dimensione intersoggettiva della proposta oscillatoria che nessun paradigma mono-soggetto può cogliere.

B.3 Analisi statistica e problemi di confronto multiplo

Le misure di PAC, PLV e TE sono statisticamente sottodeterminate quando calcolate su finestre temporali brevi — un problema particolarmente acuto nello studio del linguaggio, dove le frasi durano tipicamente 2–5 secondi e la struttura sintattica varia dinamicamente all'interno di questa finestra. I principali approcci alla validazione statistica delle misure di accoppiamento sono:

La **permutazione dei trial** — calcolare il MI su dati in cui l'associazione tra trial diversi è stata casualmente permutata, distruggendo l'accoppiamento reale ma mantenendo la struttura spettrale dei singoli segnali — fornisce la distribuzione nulla contro cui valutare la significatività del MI osservato. Il p-value è il percentile del MI osservato nella distribuzione nulla; valori $p < 0.05$ (con correzione per confronti multipli) indicano accoppiamento significativo.

Il **bootstrapping** — ricampionare i trial con reinserimento per generare una distribuzione di valori del MI sotto la variabilità campionaria — fornisce intervalli di confidenza per il MI e permette la stima della varianza inter-soggetto.

Il problema dei **confronti multipli** è particolarmente grave nelle analisi di accoppiamento che testano simultaneamente molte coppie di elettrodi e molte coppie di bande di frequenza. Le correzioni standard (Bonferroni, False Discovery Rate di Benjamini-Hochberg) sono applicabili ma spesso eccessivamente conservative. L'approccio più robusto è la correzione per permutazione cluster (*cluster-based permutation test*), che identifica cluster contigui nello spazio (elettrodi) e nella frequenza che superano una soglia primaria, e

valuta la significatività del cluster attraverso permutazioni globali — tenendo conto della correlazione spaziale e frequenziale delle misure.

B.4 Il problema della risoluzione spazio-temporale e le sue implicazioni teoriche

La tensione fondamentale nella misurazione delle oscillazioni corticali è tra la risoluzione temporale (necessaria per cogliere la dinamica dell'accoppiamento) e la risoluzione spaziale (necessaria per identificare le regioni corticali coinvolte). Questa tensione è una versione neuroscientifica del principio di incertezza: aumentare la risoluzione temporale — restringendo la finestra di analisi — riduce la risoluzione frequenziale (e dunque la capacità di distinguere le bande di frequenza), mentre aumentare la risoluzione frequenziale richiede finestre più lunghe che mascherano le variazioni temporali rapide.

Per il linguaggio, questa tensione è particolarmente acuta perché la struttura sintattica si dispiega su scale temporali che vanno dal centinaio di millisecondi (morfema, sillaba) a pochi secondi (frase, clausola) — un range di due ordini di grandezza che richiede analisi multi-scala per essere catturato adeguatamente. Le tecniche di analisi tempo-frequenza adattiva — come la trasformata wavelet di Morlet, la trasformata di Fourier a finestra mobile (STFT), o la decomposizione empirica in modi (EMD) — permettono di bilanciare risoluzione temporale e frequenziale adattivamente, con finestre più strette alle alte frequenze (dove le oscillazioni sono rapide) e finestre più ampie alle basse frequenze (dove le oscillazioni sono lente).

La trasformata wavelet di Morlet con parametro di compressione $\sigma = 7$ cicli — la scelta più comune negli studi neurolinguistici — fornisce, per una frequenza centrale f_0 , una risoluzione temporale $\sigma_t = \sigma / (2\pi \cdot f_0)$ e una risoluzione frequenziale $\sigma_f = f_0 / \sigma$. Per $f_0 = 6$ Hz (theta): $\sigma_t \approx 185$ ms, $\sigma_f \approx 0.86$ Hz. Per $f_0 = 40$ Hz (gamma): $\sigma_t \approx 28$ ms, $\sigma_f \approx 5.7$ Hz. Questa scalatura automatica assicura un numero costante di cicli ($\sigma = 7$) in tutte le bande, rendendo comparabili le misure di accoppiamento a frequenze diverse.

Bibliografia

Nota metodologica sulla bibliografia. In un'opera che attraversa campi distanti — teoria sintattica formale, neurofisiologia computazionale, filosofia del linguaggio, fenomenologia — il rischio di citare in modo impreciso o di attribuire posizioni a testi che non le contengono esattamente è reale e non trascurabile. La presente bibliografia include esclusivamente opere di cui si ha ragionevole certezza circa gli elementi bibliografici essenziali — autore, titolo, anno,

sede di pubblicazione. Dove la certezza è parziale o dove si cita una tradizione di ricerca piuttosto che un singolo testo identificato con precisione, questo è segnalato esplicitamente. La bibliografia non è esaustiva rispetto alla letteratura esistente sui temi trattati: è un repertorio delle opere che hanno contribuito direttamente all'argomentazione del saggio.

I. Teoria sintattica e biolinguistica

Chomsky, Noam (1995). *The Minimalist Program*. MIT Press, Cambridge (MA).

Chomsky, Noam (1957). *Syntactic Structures*. Mouton, The Hague. Traduzione italiana: *Le strutture della sintassi*, Laterza, Bari, 1970.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge (MA).

Fitch, W. Tecumseh; Hauser, Marc D.; Chomsky, Noam (2002). "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science*, 298(5598), 1569–1579.

Kayne, Richard S. (1994). *The Antisymmetry of Syntax*. MIT Press, Cambridge (MA). [Linguistic Inquiry Monograph 25.]

Berwick, Robert C.; Chomsky, Noam (2016). *Why Only Us: Language and Evolution*. MIT Press, Cambridge (MA).

II. Neuroscienze delle oscillazioni e del linguaggio

Buzsáki, György (2006). *Rhythms of the Brain*. Oxford University Press, Oxford.

Giraud, Anne-Lise; Poeppel, David (2012). "Cortical oscillations and speech processing: emerging computational principles and operations." *Nature Neuroscience*, 15(4), 511–517.

Poeppel, David (2003). "The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'." *Speech Communication*, 41(1–2), 245–255.

Ding, Nai; Melloni, Lucia; Zhang, Hang; Tian, Xing; Poeppel, David (2016). "Cortical tracking of hierarchical linguistic structures in connected speech." *Nature Neuroscience*, 19(1), 158–164.

Bastos, André M.; Usrey, W. Martin; Adams, Rick A.; Mangun, George R.; Fries, Pascal; Friston, Karl J. (2012). "Canonical microcircuits for predictive coding." *Neuron*, 76(4), 695–711.

Fries, Pascal (2005). "A mechanism for cognitive dynamics: neuronal communication through neuronal coherence." *Trends in Cognitive Sciences*, 9(10), 474–480.

Tort, Adriano B.L.; Komorowski, Robert; Eichenbaum, Howard; Kopell, Nancy (2010). "Measuring phase-amplitude coupling between neuronal oscillations of different frequencies." *Journal of Neurophysiology*, 104(2), 1195–1210.

Nozaradan, Sylvie; Peretz, Isabelle; Mouraux, André (2012). "Selective neuronal entrainment to the beat and meter embedded in a musical rhythm." *Journal of Neuroscience*, 32(49), 17572–17581.

Singer, Wolf (1999). "Neuronal synchrony: a versatile code for the definition of relations?" *Neuron*, 24(1), 49–65.

III. Codifica predittiva ed energia libera

Friston, Karl J. (2010). "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, 11(2), 127–138.

Friston, Karl J. (2005). "A theory of cortical responses." *Philosophical Transactions of the Royal Society B*, 360(1456), 815–836.

Rao, Rajesh P.N.; Ballard, Dana H. (1999). "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects." *Nature Neuroscience*, 2(1), 79–87.

Clark, Andy (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, Oxford.

Hohwy, Jakob (2013). *The Predictive Mind*. Oxford University Press, Oxford.

IV. Oscillazioni, memoria di lavoro e sintassi

Lisman, John E.; Jensen, Ole (2013). "The theta-gamma neural code." *Neuron*, 77(6), 1002–1016.

Jensen, Ole; Lisman, John E. (1996). "Theta/gamma networks with slow NMDA channels learn sequences and encode episodic memory: role of NMDA channels in recall." *Learning & Memory*, 3(2–3), 264–278.

Friederici, Angela D. (2011). "The brain basis of language processing: from structure to function." *Physiological Reviews*, 91(4), 1357–1392.

Friederici, Angela D. (2017). *Language in Our Brain: The Origins of a Uniquely Human Capacity*. MIT Press, Cambridge (MA).

V. Acquisizione del linguaggio e statistica distributiva

Saffran, Jenny R.; Aslin, Richard N.; Newport, Elissa L. (1996). "Statistical learning by 8-month-old infants." *Science*, 274(5294), 1926–1928.

Goswami, Usha (2011). "A temporal sampling framework for developmental dyslexia." *Trends in Cognitive Sciences*, 15(1), 3–10.

Friederici, Angela D.; Oberecker, Regine; Brauer, Jens (2012). "Neurophysiological preconditions of syntax acquisition." *Psychological Research*, 76(2), 204–211.

VI. Oscillazioni e variazione tipologica

Ding, Nai; Melloni, Lucia; Yang, Aoli; Wang, Yipu; Zhang, Wenhe; Tian, Xing; Poeppel, David (2017). "Characterizing neural entrainment to hierarchical linguistic units using electroencephalography (EEG)." *Frontiers in Human Neuroscience*, 11, 481.

Mazaheri, Ali; Jensen, Ole (2010). "Rhythmic pulsing: linking ongoing brain activity with evoked responses." *Frontiers in Human Neuroscience*, 4, 177.

VII. Modelli matematici: sistemi dinamici e sincronizzazione

Kuramoto, Yoshiki (1984). *Chemical Oscillations, Waves, and Turbulence*. Springer, Berlin.

Strogatz, Steven H. (2003). *Sync: The Emerging Science of Spontaneous Order*. Hyperion, New York.

Strogatz, Steven H. (2000). "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators." *Physica D: Nonlinear Phenomena*, 143(1–4), 1–20.

Pikovsky, Arkady; Rosenblum, Michael; Kurths, Jürgen (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, Cambridge.

VIII. Genetica e biologia evolutiva del linguaggio

Fisher, Simon E.; Vargha-Khadem, Faraneh; Watkins, Kate E.; Monaco, Anthony P.; Pembrey, Marcus E. (1998). "Localisation of a gene implicated in a severe speech and language disorder." *Nature Genetics*, 18(2), 168–170.

Enard, Wolfgang; Przeworski, Molly; Fisher, Simon E.; Lai, Cecilia S.L.; Wiebe, Victor; Kitano, Takashi; Monaco, Anthony P.; Pääbo, Svante (2002). "Molecular evolution of FOXP2, a gene involved in speech and language." *Nature*, 418(6900), 869–872.

Hauser, Marc D.; Chomsky, Noam; Fitch, W. Tecumseh (2002). [vedi sezione I]

IX. Neuroscienze comparative: canto degli uccelli

Hahnloser, Richard H.R.; Kozhevnikov, Alexay A.; Fee, Michale S. (2002). "An ultra-sparse code underlies the generation of neural sequences in a songbird." *Nature*, 419(6902), 65–70.

Long, Michael A.; Fee, Michale S. (2008). "Using temperature to analyse temporal dynamics in the songbird motor pathway." *Nature*, 456(7219), 189–194.

X. Patologie del linguaggio

Friederici, Angela D. (2002). "Towards a neural basis for auditory sentence processing." *Trends in Cognitive Sciences*, 6(2), 78–84.

Goswami, Usha (2011). [vedi sezione V]

Friston, Karl J.; Lawson, Rebecca (2012). "Explaining Away and Prediction Error: Response to the Commentaries." *Psychological Medicine*, 42(5), 1153–1156. [Per il modello dell'autismo come iper-precisione sensoriale.]

XI. Musica e linguaggio: struttura e neuroscienze

Koelsch, Stefan (2012). *Brain and Music*. Wiley-Blackwell, Oxford.

Koelsch, Stefan; Gunter, Thomas C.; Friederici, Angela D.; Schröger, Erich (2000). "Brain indices of music processing: 'non-musicians' are musical." *Journal of Cognitive Neuroscience*, 12(3), 520–541.

Hanslick, Eduard (1854). *Vom Musikalisch-Schönen: Ein Beitrag zur Revision der Ästhetik der Tonkunst*. Rudolph Weigel, Leipzig. Traduzione italiana: *Il bello musicale*, Aesthetica, Palermo, 2001.

Patel, Aniruddh D. (2008). *Music, Language, and the Brain*. Oxford University Press, Oxford.

XII. Coscienza, global workspace e teoria dell'informazione integrata

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press, Cambridge.

Dehaene, Stanislas; Changeux, Jean-Pierre (2011). "Experimental and theoretical approaches to conscious processing." *Neuron*, 70(2), 200–227.

Tononi, Giulio (2004). "An information integration theory of consciousness." *BMC Neuroscience*, 5, 42.

Tononi, Giulio; Koch, Christof (2015). "Consciousness: here, there and everywhere?" *Philosophical Transactions of the Royal Society B*, 370(1668), 20140167.

Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, Oxford.

Metzinger, Thomas (2003). *Being No One: The Self-Model Theory of Subjectivity*. MIT Press, Cambridge (MA).

XIII. Filosofia del linguaggio e fenomenologia

Austin, John L. (1962). *How to Do Things with Words*. Oxford University Press, Oxford. Traduzione italiana: *Come fare cose con le parole*, Marietti, Genova, 1987.

Searle, John R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge.

Grice, Herbert Paul (1975). "Logic and Conversation." In: Cole, Peter; Morgan, Jerry L. (a cura di), *Syntax and Semantics*, vol. 3: *Speech Acts*. Academic Press, New York, pp. 41–58.

Wittgenstein, Ludwig (1953). *Philosophische Untersuchungen / Philosophical Investigations*. Blackwell, Oxford. Traduzione italiana: *Ricerche filosofiche*, Einaudi, Torino, 1967.

Husserl, Edmund (1928). *Vorlesungen zur Phänomenologie des inneren Zeitbewußtseins*. Halle. Edizione critica: Husserliana X, Nijhoff, Den Haag, 1966. Traduzione italiana: *Per la fenomenologia della coscienza interna del tempo*, Franco Angeli, Milano, 1981.

Merleau-Ponty, Maurice (1945). *Phénoménologie de la perception*. Gallimard, Paris. Traduzione italiana: *Fenomenologia della percezione*, Il Saggiatore, Milano, 1965.

Frege, Gottlob (1892). "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50. Traduzione italiana: "Senso e denotazione", in: *Logica e aritmetica*, Boringhieri, Torino, 1965.

XIV. Linguaggio interno e pensiero

Vygotsky, Lev S. (1934). *Мышление и речь*. Sotsekgiz, Moskva-Leningrad. Traduzione italiana: *Pensiero e linguaggio*, Laterza, Bari, 1992.

Fodor, Jerry A. (1975). *The Language of Thought*. Harvard University Press, Cambridge (MA).

XV. Bilinguismo e code-switching

Myers-Scotton, Carol (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Clarendon Press, Oxford.

Grosjean, François (1982). *Life with Two Languages: An Introduction to Bilingualism*. Harvard University Press, Cambridge (MA).

XVI. Sonno e consolidamento mnemonico

Stickgold, Robert (2005). "Sleep-dependent memory consolidation." *Nature*, 437(7063), 1272–1278.

Diekelmann, Susanne; Born, Jan (2010). "The memory function of sleep." *Nature Reviews Neuroscience*, 11(2), 114–126.

XVII. Segnali di errore e componenti ERP

Kutas, Marta; Federmeier, Kara D. (2011). "Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP)." *Annual Review of Psychology*, 62, 621–647.

Friederici, Angela D.; Pfeifer, Elisabeth; Hahne, Anja (1993). "Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations." *Cognitive Brain Research*, 1(3), 183–192.

XVIII. Iper-scanning e sincronizzazione inter-cerebrale

Hasson, Uri; Ghazanfar, Asif A.; Galantucci, Bruno; Garrod, Simon; Keysers, Christian (2012). "Brain-to-brain coupling: a mechanism for creating and sharing a social world." *Trends in Cognitive Sciences*, 16(2), 114–121.

Nota sulle opere non citate per cautela

Diverse opere e studi specifici menzionati nel corso del saggio — in particolare alcuni articoli di EEG/MEG su popolazioni specifiche (sordi segnanti, bilingui tardivi, pazienti afasici), studi longitudinali sullo sviluppo, e lavori di modellazione computazionale sulle oscillazioni sintattiche — non compaiono in bibliografia perché non è stato possibile verificarne con sufficiente certezza gli estremi bibliografici precisi. Per questi ambiti, si rimanda alle rassegne sistematiche citate — in particolare Friederici (2017), Buzsáki (2006), e Patel (2008) — che forniscono un accesso strutturato alla letteratura primaria rilevante, con bibliografia secondaria accurata e aggiornata.